

**PHYSICAL SIGNAL-BASED INTRUSION DETECTION FOR CYBER  
PHYSICAL SYSTEMS**

A Thesis  
Presented to  
The Academic Faculty

By

Christian J. Bayens

In Partial Fulfillment  
of the Requirements for the Degree  
Master's of Science in the  
School of Electrical and Computer Engineering

Georgia Institute of Technology

May 2017

Copyright © Christian J. Bayens 2017

**PHYSICAL SIGNAL-BASED INTRUSION DETECTION FOR CYBER  
PHYSICAL SYSTEMS**

Approved by:

Dr. Beyah, Advisor  
School of Electrical and Computer  
Engineering  
*Georgia Institute of Technology*

Dr. Cohen  
School of Electrical and Computer  
Engineering  
*Georgia Institute of Technology*

Dr. Zonouz  
School of Electrical and Computer  
Engineering  
*Rutgers University*

Date Approved: April 25, 2017

I don't know anything, but I do know that everything is interesting if you go into it deeply  
enough.

*Richard Feynman*

## **ACKNOWLEDGEMENTS**

I would like to thank my advisor, Dr. Beyah, for his guidance and enthusiasm throughout this research. I would also like to thank Dr. Cohen and Dr. Zonouz for the significant roles that they played in guiding and forming the research.

I would like to also thank Jackson McCormick for his technical assistance as well as Dave Robinson for his help in data collection. Finally, thank you to all the family, friends, faculty and staff that have helped along the way.



## TABLE OF CONTENTS

<b>Acknowledgments</b> . . . . .	v
<b>List of Figures</b> . . . . .	viii
<b>Chapter 1: Introduction</b> . . . . .	1
1.1 Additive Manufacturing . . . . .	1
1.2 Electrical Grid . . . . .	3
1.3 Contributions and Thesis Organization . . . . .	4
<b>Chapter 2: Related Work</b> . . . . .	6
2.1 Threats to Cyber-Physical Systems . . . . .	6
2.2 Acoustic, Magnetic, and Motion Sensing . . . . .	7
2.3 RF Monitoring of Power Systems . . . . .	7
2.4 AWESOME Receiver . . . . .	8
2.5 National Lightning Detection Network . . . . .	8
<b>Chapter 3: Threat Model</b> . . . . .	9
<b>Chapter 4: Additive Manufacturing Intrusion Detection</b> . . . . .	10
4.1 Intrusion Detection Methods . . . . .	10
4.2 Evaluation . . . . .	16

4.2.1	Classification Accuracy . . . . .	16
4.2.2	Varied Printer Models . . . . .	20
4.2.3	Detecting Extrusion . . . . .	21
4.2.4	Classification with Minimal Change . . . . .	22
4.2.5	Classification in Noisy Environments . . . . .	23
4.2.6	Visualization of Malicious Prints . . . . .	23
4.2.7	Case Study: Prosthetic Knee . . . . .	24
4.3	Implementation . . . . .	27
<b>Chapter 5: Electrical Grid Intrusion Detection . . . . .</b>		<b>29</b>
5.1	Switching Detection Methods . . . . .	29
5.2	Evaluation . . . . .	32
5.2.1	Network Traffic Comparison . . . . .	34
5.2.2	Electrical Grid Event Detection at a Distance . . . . .	35
5.3	Implementation . . . . .	37
<b>Chapter 6: Conclusion and Future Work . . . . .</b>		<b>39</b>
6.1	Conclusion . . . . .	39
6.2	Future Work . . . . .	39
<b>Chapter A: Detailed Results of Acoustic Classification on Tibial Knee Prosthetic</b>		<b>42</b>
<b>References . . . . .</b>		<b>45</b>

## LIST OF FIGURES

1.1	High level view of physical signal-based intrusion detection. . . . .	4
4.1	Diagram of audio classification model. . . . .	11
4.2	Top Hat and Rectangular Prism designs. . . . .	12
4.3	Classification example. . . . .	14
4.4	Spatial sensing setup with Unimeasure linear potentiometer model number LA-PA-10-N1N-NPC, SparkFun Triple Axis Accelerometer and Gyro Breakout, and Teensy 3.2 board. . . . .	14
4.5	Comparison of G-code reconstruction to gyroscopic sensing. . . . .	15
4.6	ROC curves for Rectangular Prism, CTh = 35. . . . .	17
4.7	ROC curves for Top Hat. . . . .	19
4.8	Comparison of the frequency response between a single layer of Honeycomb 40% fill and Rectilinear 40% fill. . . . .	20
4.9	ROC curves for top hat design printed using a TazMini, Orion Delta, and Taz6 print. Print audio sliced to 120 seconds and the confidence threshold is 150, 20, and 35 respectively. . . . .	21
4.10	Calculated scores differentiating prints that extrude material and those that do not. . . . .	22
4.11	ROC curve for malicious print detection between Honeycomb and Rectilinear prints with 20%, 30% and 40% fill density. . . . .	23
4.12	Comparison of acoustic print classification for target tibial prosthetic with 60% Rectilinear Fill (Top) vs. malicious 20% Honeycomb Fill (bottom). CTh = 0. . . . .	25

4.13	Comparison of x-axis frequency response for a layer of the tibial knee implant design. . . . .	26
4.14	Comparison of benign and malicious tibial knee implant prints. . . . .	26
4.15	Classification data from attempted physical replay attack. . . . .	28
5.1	Diagram of a simple electrical grid [31] . . . . .	29
5.2	AWESOME reciever and antenna setup. . . . .	30
5.3	Sample recorded LF data with two groups of lightning strikes. . . . .	31
5.4	Current C1 opened. . . . .	32
5.5	Current C2 opened. . . . .	32
5.6	Current C1 closed. . . . .	33
5.7	Current C3 opened. . . . .	34
5.8	Current C4 opened. . . . .	34
5.9	Network traffic for currents C1 and C2 opened. . . . .	35
5.10	Distance from Sub2 to AWESOME network receiver. . . . .	36
5.11	Transformer switch. . . . .	36
5.12	Transformer switch detail. . . . .	36
5.13	Distance from Sub3 to AWESOME network receiver. . . . .	37
5.14	AWESOME data during Sub3 outage. . . . .	37

## SUMMARY

In recent years, a significant emerging target of cyber-based attacks by nation-states and other advanced groups is cyber-physical systems (CPS). These attacks target major utility, manufacturing, or public service infrastructure in order to collect ransom or intellectual property. A major feature of these attacks is the ability to spoof network traffic to indicate normal activity to a user while malicious instructions are sent to the physical machinery.

This thesis investigates methods by which physical sensing and signal analysis may be used as intrusion detection in such a scenario. For the sector of additive manufacturing (AM), we use audio classification and motion detection to identify malicious prints. For the power utility sector, we use a specialized low-frequency radio receiver to detect switching events which can then be compared to network traffic to detect the presence of malicious activity.

# **CHAPTER 1**

## **INTRODUCTION**

Cyber-physical systems (CPS) are, simply put, any mechanical system that is integrated with computer algorithms and the Internet for monitoring and control. These may include infrastructure for manufacturing, utilities, HVAC, etc. Evidence from both world events such as the infamous STUXNET and Ukrainian power grid attacks as well as recent academic research suggest that attacks on cyber-physical systems are on the rise [1, 2, 3]. These attacks generally focus on bypassing network-based intrusion detection systems and setting up a man-in-the-middle attack so that malicious activity can be performed without evidence on the human-machine interface (HMI). Therefore, this thesis focuses on the analysis of physical signals caused by the machinery being controlled by the CPS and its usefulness for intrusion detection. The types of CPS infrastructure explored are Additive Manufacturing (AM) and the electrical grid infrastructure.

### **1.1 Additive Manufacturing**

Additive Manufacturing (AM), also known as 3D printing, is an emerging field that shows promise in reducing waste, time, and infrastructure needed in a manufacturing process. Many major companies including Ford, GE, Airbus, SpaceX, Koenigsegg, and NASA are currently utilizing AM for both prototyping and production-quality manufacturing [4, 5, 6, 7, 8, 9]. Additionally, AM has been employed as a useful tool for printing medical implants [10], and cutting edge research is underway on producing food, drugs, and living tissue using AM techniques [11, 12]. Across industries, AM is expected to reach a market potential of 50% by 2038 [13].

Because of this potential for wide-spread use of AM in the coming decades, work has begun on understanding the security challenges that uniquely differ from traditional man-

ufacturing and cyber-physical security. Mark Yampolskiy, *et al.*, [14] outlined a taxonomy for the potential of the misuse of a 3D printer as a weapon (3D-PaaW). In their thesis, they identify the elements which may compromise or manipulate an AM environment, the targets of attack (printed object, printers, or environment), and the parameters for understanding the potential effectiveness of a given attack.

In this thesis, we focus on the use of a 3D-PaaW to manipulate the physical properties of a printed object through manipulation of the object specifications, manufacturing parameters, and/or source material. According to the taxonomy described by Yampolskiy, *et al.*, each of these are classified as attacks which would be achievable by an adversary through the manipulation of printer firmware. It has been shown that structural integrity can be easily compromised by introducing slight modifications in the model, e.g., a minuscule void injected into a manufactured object can reduce the yield load by 14 percent [15].

In order to combat these forms of attack, we propose two methods of verification of design parameters that utilize analysis of the acoustic signal and spatial position of machine components. Intrusion detection is then achieved in the event that a malicious print fails this verification process. Acoustic and spatial methods are chosen because they provide information about the manufactured design *without* the need to reference the STL file or the G-code instructions<sup>1</sup> read by the printer. We do not consider our techniques to be a panacea for all verification needs. They are meant to be complementary to domain-specific verification methods. In some cases, this may be means of saving costs, e.g., by detecting malicious prints in real-time and ending them at the onset of a detection. In other cases, this may be a means of ensuring safety, e.g., by detecting malicious materials or designs before the print is used.

---

<sup>1</sup>An STL file is a STereoLithography file for CAD software used in 3D printing. G-code is the set of actual instructions for 3D printers that are generated for particular models given an STL file and the print configuration, e.g., print speed and infill density.

## 1.2 Electrical Grid

While AM is an emerging industry, the power grid is a well established and , in some cases, aging infrastructure. Due to increasing demand, the need for greater efficiency, and a number of other factors, significant work has been undertaken in the development of the "Smart Grid". Smart grid technology uses supervisory control and data acquisition (SCADA) systems in order to control and monitor widespread power systems [16]. Unfortunately, the in-depth control and supervision that is made possible by SCADA also results in the possibility of cyber attacks.

One such example is that of that well known attack on the Ukrainian power grid. The attack used a technique called spear-phishing in order to gain access to the smart grid network by posing as a trusted entity [17]. The result was 30 substations being switched off and more than 230,000 residents left without power. A key aspect of the attack was a distributed denial of service (DDOS) attack on the call centers so that customer complaints could not be made to the power company. Between this and the spoofing of network traffic, the company was unaware of the attack until it was too late. At this point, the substations were shut off and would not accept commands from the power company to come back online.

This thesis explores methods of analyzing the low frequency (LF) electromagnetic signals generated by the 60Hz alternating current of the electrical grid. In theory, when the substations involved in the Ukrainian power grid attack switched off, the sudden loss of current should have caused a brief spike in the magnetic field. Despite the cleverness of the hackers, this is an effect that cannot be physically faked. Therefore, if this magnetic field spike were able to be detected, the power company would not have had to rely on blocked complaint calls from its customers to identify the network intrusion.



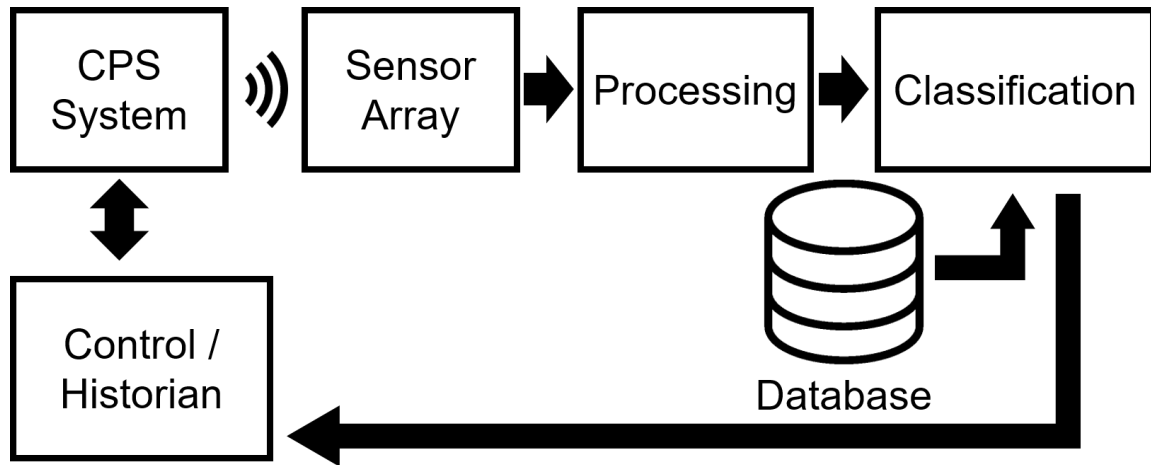


Figure 1.1: High level view of physical signal-based intrusion detection.

### 1.3 Contributions and Thesis Organization

The primary contributions of this thesis are as follows:

- Audio classification methods for intrusion detection in additive manufacturing systems
- LF magnetic field analysis methods for intrusion detection in electrical grid systems.
- Suggested implementation methods for physical signal analysis-based intrusion detection.

A generalized high-level view of the proposed physical signal-based intrusion detection system is shown in Figure 1.1. Physical signals are detected by an array of sensors, processed using various signal processing techniques, and then classified and compared to data previously stored in a database. The information gained from this classification is then compared to data in the control system and/or historian.

The remainder of the thesis is organized as follows. Chapter 2 presents related work in the areas of 3D printing security, audio classification, electrical grid security, LF radio analysis, and lightning detection. Chapter 3 outlines a generalized threat model. Chapter

4 discusses the methods for intrusion detection in AM and evaluates the methods. Chapter 5 discusses and evaluates the methods for electrical grid intrusion detection. Chapter 6 discusses implementation of physical signal-based intrusion detection. Finally, Chapter 7 concludes and discusses future work.

## **CHAPTER 2**

### **RELATED WORK**

In this thesis we provide a means of intrusion detection by utilizing the various signals and side-channels associated with physical processes. As such, we first review previous efforts that have been made for the analysis of the side-channels in the AM process. We then briefly describe the previous work that allows for the highly accurate detection of low frequency radio signals.

#### **2.1 Threats to Cyber-Physical Systems**

In recent years, widely publicized attacks such as STUXNET [1] and the attack on the Ukrainian power grid [17] have shown the capability of attackers to attack cyber-physical networks. These attacks are characterized by the ability of an attacker to manipulate the physical end of the system while feigning normalcy to the user.

Formby, *et al.*, showed how a water treatment plant can be the target of a ransomware attack. In the attack, sensor data is spoofed such that the plant is unaware of the attack until it is made aware by the attacker. At this point, the plant must either pay the ransom or else the attacker raises the chlorine levels to a dangerously high level [2]. Likewise, Garcia, *et al.*, developed a root kit that allows for physics-aware sensor spoofing during an attack [18]. Finally, Fachkha, *et al.*, have shown significant numbers of organizations or individuals specifically probing CPS-related IP space, likely in reconnaissance for future attempted attacks [3].

This thesis aims to develop monitoring techniques that are very difficult or impossible to spoof in order to provide immediate detection of intrusion.

## 2.2 Acoustic, Magnetic, and Motion Sensing

KCAD [19] provided the first method of using the analog emissions of AM processes for the purpose of detecting so-called zero-day kinetic cyber-attacks. However, the work utilizes only one 3D printer and only investigates attacks that inject simple variations in the exterior design.

The focus of the majority of previous work on the analysis of side-channels from 3D printers used in AM has been its usefulness in obtaining intellectual property. Chen Song, *et al.*, [20] and Avesta Hojjati, *et al.*, [21] each showed that the array of sensors available on a modern smart phone can be leveraged to re-create designs produced from 3D printers or CNC machines. The sensors used in each study to collect side-channel data included the microphone, magnetometer, and accelerometer. Each group was able to reconstruct simple printed designs using supervised machine learning and manual analysis of sensor signals respectively. However, each group was only able to reconstruct very simple shapes such as two-dimensional outlines of airplanes or keys with no fill structure.

## 2.3 RF Monitoring of Power Systems

Most of the previous work in the monitoring of power stations using RF has focused on issues of power quality. Baker, *et al.*, investigate the use of RF to detect partial discharge which can deteriorate power system equipment [22]. Likewise, Nesbitt, *et al.*, assessed the condition of high voltage systems using RF [23]. Each of these systems relied on frequencies in the megahertz range. In the case of Baker, *et al.*, , this resulted in the need to directly attach the sensor to the equipment it was measuring to minimize signal attenuation. Also Nesbitt, *et al.*, , found they needed to constantly tune the receiver to obtain useful data.

In this thesis, we use a system that measures low frequency ( $< 500\text{kHz}$ ) signals which are within range of the RF emissions generated directly by power lines.

## 2.4 AWESOME Receiver

In 2010, Cohen, *et al.*, developed an instrument called Atmospheric Weather Electromagnetic System for Observation, Modeling, and Education (AWESOME) [24]. AWESOME was designed as a scientifically sensitive radio receiver that could measure broadband frequencies in the ELF/VLF spectrum (300Hz - 30kHz). The system has more recently been upgraded to cover the frequency band up to 500 kHz. The primary use for this instrument is to measure atmospherics (sferics) which are ELF/VLF signals reflected by the D-region of the ionosphere. Sferics are most commonly generated naturally by lightning or by radiation effects in the magnetosphere [24].

Cohen, *et al.*, identify power lines as a significant source of interference when detecting atmospherics. This is due to the strong magnetic field caused by the 50/60Hz current which produces harmonics up to several kilohertz. In atmospheric research, it is useful to avoid such signals or to filter them out [25]. However, in this thesis, the goal is to directly measure the signals from power lines in order to detect power line activity.

## 2.5 National Lightning Detection Network

In measuring power system activity using the AWESOME receiver, it is important to be able to differentiate these man-made signals from natural signals such as those caused by lightning. To do this we use the National Lightning Detection Network (NLDN) [26].

NLDN is widely used amongst government agencies, scientific groups, and weather monitor for accurate tracking of the location, time, intensity, and type of lightning strike throughout the 48 contiguous United States. Its specifications include accuracy location accuracy of  $< 200\text{m}$ , cloud-to-ground and cloud-to-cloud lightning detection efficiency of 95% and 50-60% respectively, and Event timing precision of 0.5 microsecond RMS.

## **CHAPTER 3**

### **THREAT MODEL**

Intrusion detection systems must be designed with specific types of intrusions in mind. With that, we explicitly specify the threat model that this system is designed to defend against.

The threat model assumes that the attacker has full knowledge of both the physical system and its control software. Any third party manufacturer or affiliate of the user is trusted as an organization. Therefore, they are willing to provide information about the print for verification. However, malicious entities may include network intruders with access gained via phishing, social engineering, or other common methods. The attack is carried out such that the system behaves maliciously despite being sent benign instructions.

In AM, the printer behaves maliciously despite being sent G-code for a benign print. The malicious firmware may contain a hard-coded design that would replace the one sent by the user or otherwise manipulate otherwise benign commands. This has been shown to be feasible through manipulation of printer firmware [27].

For the electrical grid, an attacker is capable of sending commands on the network and manipulating traffic to the user in a similar style as the STUXNET attack [1]. In this scenario, either a circuit does not open or close despite the proper command being sent, or conversely, the circuit opens or closes with no apparent command. Meanwhile, the HMI indicates that commands or lack thereof are carried out normally. This attack is feasible using a cyber-physical rootkit such as the one described by Garcia, *et al.*, [18] or by methods described by Formby, *et al.*, [2].

## CHAPTER 4

### ADDITIVE MANUFACTURING INTRUSION DETECTION

#### 4.1 Intrusion Detection Methods

**Acoustic Classification.** As a physical byproduct of nearly any mechanical process, acoustic signals have been explored as a method of understanding information being processed by both traditional printers [28] and 3D Printers used in AM [20, 21, 19]. Because traditional printing methods now rely on lasers or ink jets, the information obtained from these is minimal. However, 3D printers will continue to rely on various actuators and fans for the foreseeable future which produce useful acoustic data. This is especially true for large-scale implementations of the technology.

For intrusion detection through audio classification, we assume that a particular design with a given infill structure will be printed multiple times. We use an open source audio classifier similar to the Shazaam [29] or SoundHound Applications. Using a training audio file, it locates noise-resistant peak frequencies and their temporal location within the file. Each tuple of frequency and time are then hashed using SHA1<sup>1</sup> and classification is performed by searching for collisions. When a test file is classified, it is accompanied by a confidence score among other information. The confidence score indicates the number of peaks that the test has in common with the training data.

For AM verification, we use a single print as a training set by recording it with a microphone to obtain an audio file. Because even a simple print can take several minutes, the resulting file is separated into a number of segments of a given length (some number of seconds) and indexed in ascending order. Each indexed segment of the print is then trained as a unique audio file and stored in a database. The process is illustrated in Figure 4.1

---

<sup>1</sup>SHA1 is known to be unsecure, but security of the hash itself here is not necessary

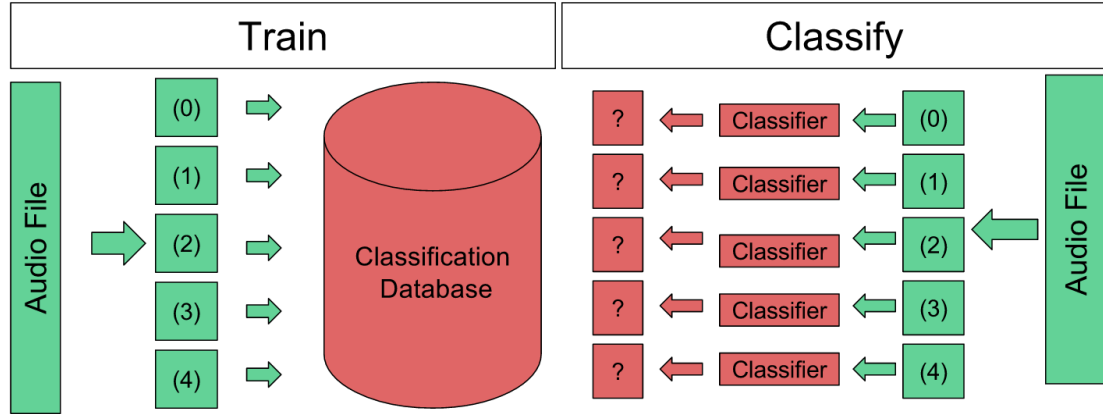


Figure 4.1: Diagram of audio classification model.

In many machine learning schema, common practice is to train on multiple sets of data. However, because acoustic classification involves one-to-one comparison of audio files, a single-file training set is appropriate.

Identical G-Code is used to produce identical prints. Identical prints may be achieved through non-identical G-Code by using varied slicing algorithms, but that is beyond the scope of this thesis.

Test data is collected using the same method as training data and split into segments of the same length. Each indexed segment is then classified independently and a confidence score is returned. The confidence score represents the number of frequency peaks that a given file has in common with the training file. Verification that a test print is unaltered from the training print is determined in two ways:

1. The classification results are such that the index values appear in the same order that they are classified in. For example if the indexes 1, 2, and 3 of the test print return classifications 2, 3, and 1 respectively, then the verification has failed.
2. The confidence score of one or more indexed classification results falls below a given threshold value. The threshold value is referred to as the confidence threshold (CTh) for the remainder of the thesis. Its value is optimized for each individual printer to



maximize the true positive rate and minimize the false positive rate.

With this, a print will be considered verified if each indexed audio file is classified correctly, in the correct order, and with confidence values greater than CTh. A non-verified print, conversely, will be classified out of order and/or with one or more confidence values less than CTh.

To test this method, two designs, shown in Figure 4.2, are used throughout this section. They are described as a Rectangular Prism (right) and a Top Hat (left). Each was printed several times with “Honeycomb” and “Rectilinear” fill patterns of 20%, 40%, and 60% density. For each print style, a single set of audio data was split and stored in a unique database as described above.

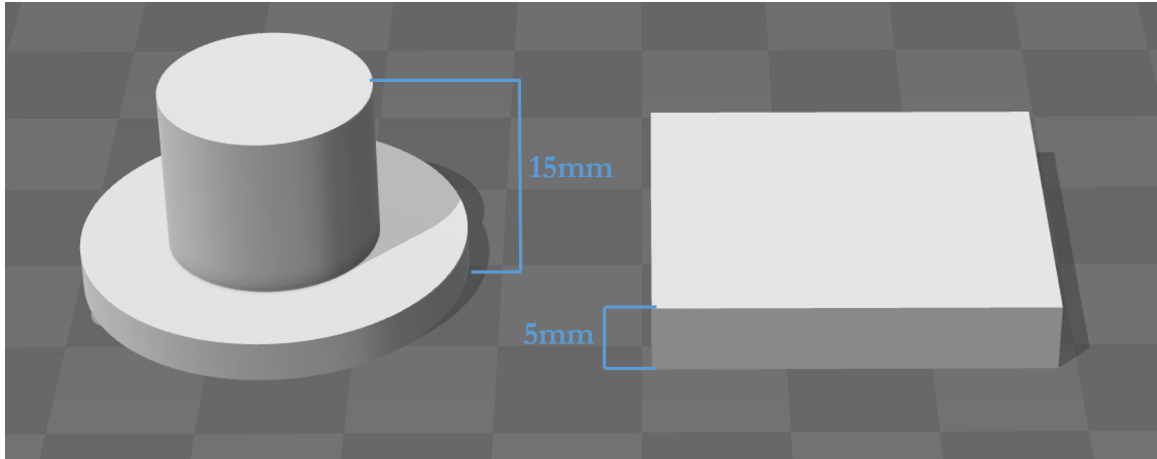


Figure 4.2: Top Hat and Rectangular Prism designs.

In order to derive quantitative results to the test classifications, we assign a “score” to each segment of the audio data which are defined as follows:

- If a segment is in proper sequence and the confidence value is greater than CTh, its score is equal to that of the confidence value.
- If a segment is out of sequence, its score is equal to  $-1 * \text{confidence value}$ .
- If a segment is in sequence, but the confidence value is less than CTh, its score is set equal to  $-1 * \text{confidence value}$ .

If a very small or negative score is calculated for any segment of the sliced audio file, a positive intrusion classification may be determined. If no negative values are calculated, a negative intrusion classification is determined.

Sample results are shown in Figure 4.3. The print is a Rectangular Prism with a 20% density Honeycomb fill pattern. The top chart shows the averaged results of three known benign prints with no malicious changes (true negatives). Each bar represents a 90 second slice of the printing data, and CTh is set to 35. Likewise, the bottom chart represents various malicious prints (true positives) caused by incorrect fill densities or patterns. Each type of error is printed four times and the results are averaged. For malicious prints with a Honeycomb fill at 60% density, a positive intrusion classification is achieved within 270s or the first 60% of the print. This is shown by the calculated score becoming very small between 181 and 210 seconds into the print. For the erroneous Rectilinear fill patterns and the Honeycomb fill with 40% density, positive error classification is achieved within 180s or 40% of the print. Again, this is seen by the calculated score becoming very small or negative between 91 and 180 seconds into the print. In each case, the first 90 seconds of the malicious prints always receive high scores due to the fact that the design always starts with a 100% rectilinear fill of the first three layers. This is standard in 3D printing to ensure that the exterior is solid.

**Spatial Sensing.** When performing 3D prints, it was found that the software used to monitor print progress simply displayed the progress of the G-code instructions being sent to the printer. This is regardless of the actual actions of the printer. The goal in setting up a spatial sensing verification scheme was to physically monitor the position of the printing nozzle with respect to the printing base, in order to observe their actual positions throughout the printing process.

The first consideration was to use a ride-along accelerator. However, due to the double integration from acceleration to position and the noisiness of the accelerometer data, visual

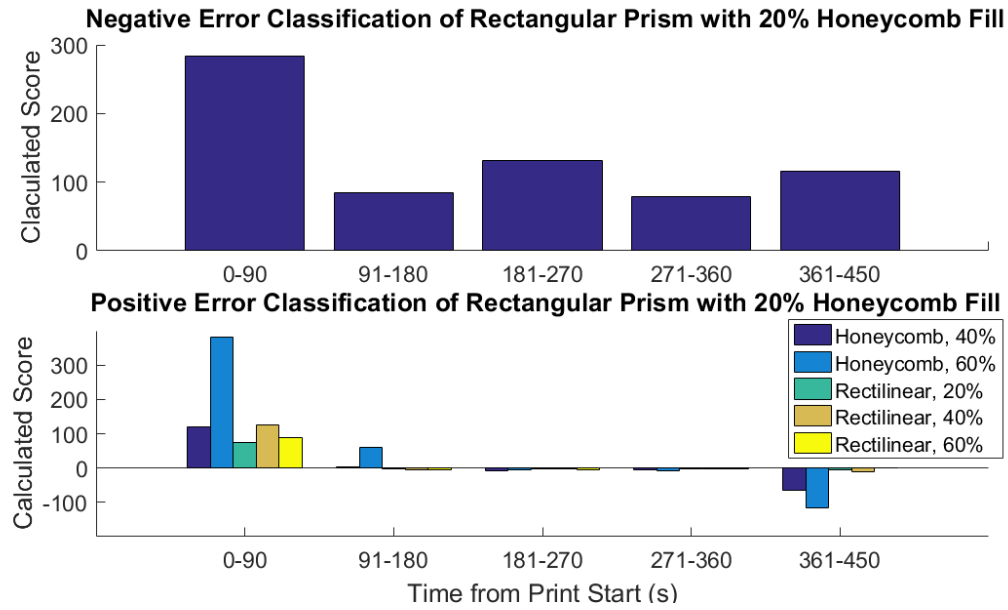


Figure 4.3: Classification example.

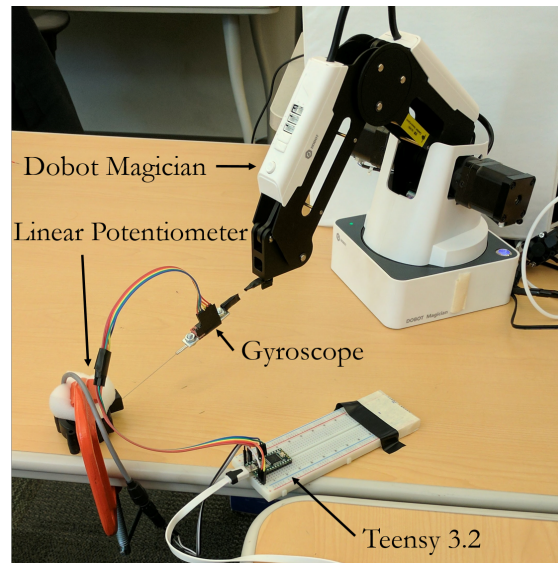


Figure 4.4: Spatial sensing setup with Unimeasure linear potentiometer model number LA-PA-10-N1N-NPC, SparkFun Triple Axis Accelerometer and Gyro Breakout, and Teensy 3.2 board.

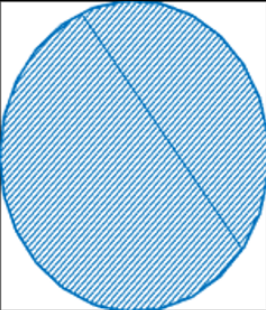
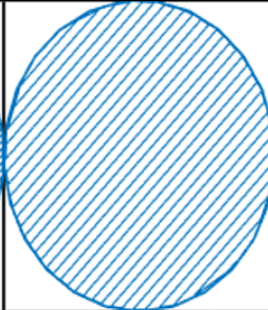
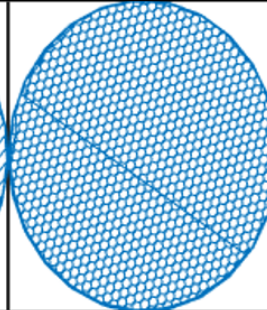
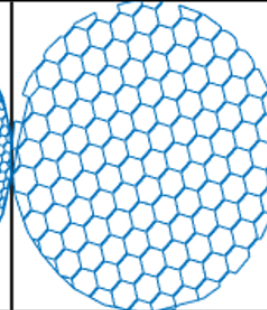
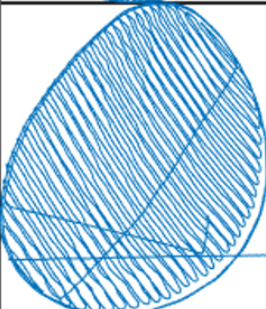
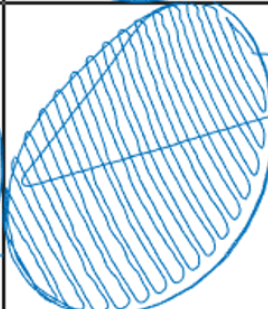
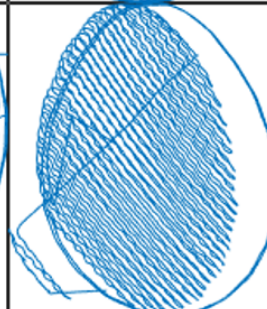
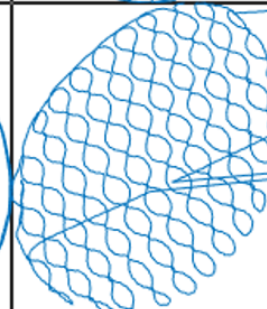
Rectilinear 60%	Rectilinear 20%	Honeycomb 60%	Honeycomb 20%	
				G-Code  Spatial Reconstruction
				

Figure 4.5: Comparison of G-code reconstruction to gyroscopic sensing.

representations of the printer’s path became prohibitively difficult to obtain. An attempt was also made to gather data from a 3D printer that used a built-in accelerometer for calibration purposes. However, the data collected from it was very noisy and low resolution.

A scheme was then developed in which a gyroscopic sensor was paired with a linear potentiometer in order to construct a set of spherical coordinates to describe the printer’s motion. This proved more effective because no integration was needed for the data, and only simple moving average filtering was necessary to reduce noise to a usable level.

To obtain these measurements, the following devices were used: a Unimeasure linear potentiometer model number LA-PA-10-N1N-NPC, a SparkFun Triple Axis Accelerometer and Gyro Breakout MPU-6050, and a Teensy 3.2 board. The experiments were conducted in a setup as shown in Figure 4.4 with a Dobot Magician desktop CNC and 3D Printer. For experimental purposes, the actual 3D printing extruder was removed and “dummy” prints were performed. The test prints were a single layer of a circular disk printed with Honeycomb and Rectangular fills each with a 20% and 60% density. Data is

collected at a rate of 100Hz. In Figure 4.5, each print is shown as the G-code representation next to the reconstructed path of the printer. The data shown is smoothed using a moving average filter with a window of five. This causes a rounding effect in the Honeycomb features, but removes random noise that makes the structures difficult to see.

## 4.2 Evaluation

In this section, we evaluate the usefulness of the proposed intrusion detection method for AM. This initial identification will be carried out primarily by acoustic classification with redundancy in the spatial sensing to reduce false classifications.

### 4.2.1 Classification Accuracy

In order to gain initial understanding of the parameters that affect the accuracy classification for intrusion detection, several experiments were carried out with a small number of trials. The printers used in the tests described in this section were a Lulzbot Taz6, Lulzbot TazMini, an Orion Delta. The AKG P170 condenser microphone was placed on a stand as close to the moving extruder head without being knocked over by the moving components of the printer. The software used for audio classification is called dejavu [30] and is an open-source project written in python.

In order to generate data useful for logistic regression, a vector of scores,  $\mathbf{S}$ , is generated using the exact method as is described in section 4.1. For example, the components of  $\mathbf{S}$  are shown in Figure 4.3. The vector  $\mathbf{S}$  is of length  $n$  where  $n = \lfloor \frac{\text{audio length}}{\text{audio slice length}} \rfloor$ . We then calculate a print score,  $p$ , where

$$p = \sum_n S_n . \quad (4.1)$$

The value  $p$  associated with a given print now determines how likely the print is to be the same as the training print. Higher values indicate more likely and lower values indicate

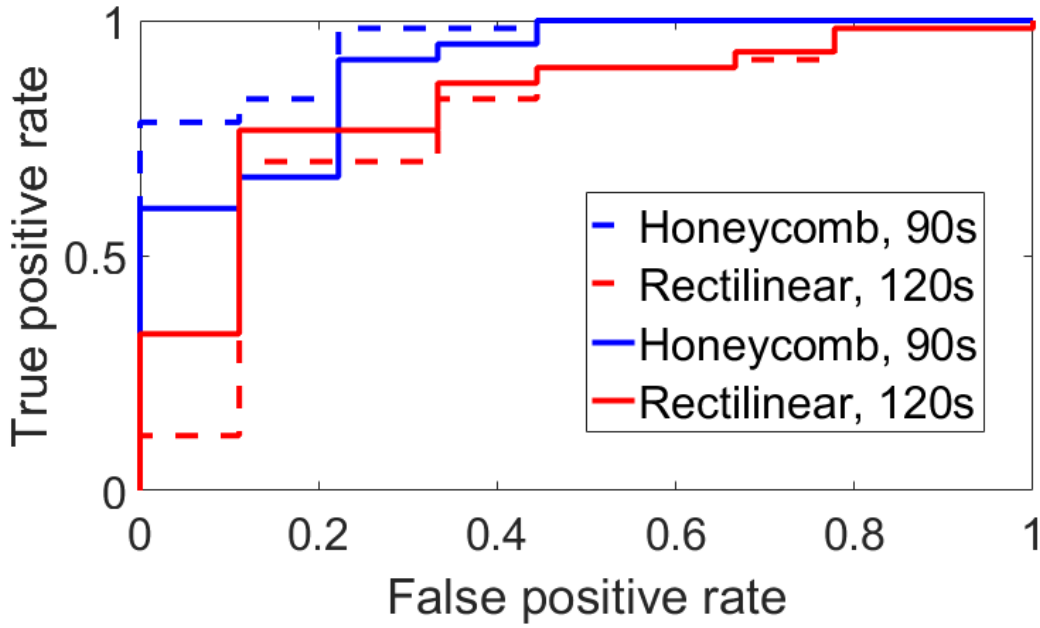


Figure 4.6: ROC curves for Rectangular Prism,  $C_{Th} = 35$ .

less likely.

To quantify the accuracy of the results of the various tests, the data is fit into a logistic regression model with the binary dependent variable of “malicious print detected” or “no malicious print detected”. From the model, we extract the probabilistic classification outcomes and create a receiver operating characteristic (ROC) curve. The area under the ROC curve (AUROC) is the metric used to define classification accuracy.

In Figure 4.6, the ROC curves are shown for the classification results a Rectangular Prism (Figure 4.2) with Honeycomb and Rectilinear fills. The audio is segmented to 90 second and 120 second segments. For each,  $C_{Th}$  was set equal to 35 as the integer value that maximized true positives without increasing false positives. The same original audio files are used whether the audio files are segmented to 90 seconds or 120 seconds. The Honeycomb and Rectilinear tests each consist of nine benign prints and sixty malicious prints. The reason for the mismatch of benign and malicious print is that each print is considered benign when compared to a copy of itself, but malicious when compared to a

different design.

The poorest performance shows an AUROC value of 0.7815 for the rectilinear fill with the audio segmented at 90 seconds. That was determined to be unacceptable especially considering the high likelihood of false positives. To find an explanation for the poor classification, the G-code was inspected. Upon investigation of the G-code which was generated by Slic3r, a group of 9 lines were identified which repeated 12 times each out of the 15 layers needed to complete the print in both the Rectilinear and Honeycomb fill patterns. These lines specified  $x$  and  $y$  coordinates along with the extrusion rate. Also, upon investigating sequentially repeated blocks of code, it was found that blocks of G-code describing three entire layers were repeated twice during the course of the print. This high repetition and symmetry was hypothesized to be the cause of the classification confusion because the audio from these identical segments may be misclassified as each other.

To test this hypothesis, a second set of tests were conducted with a Top Hat design (Figure 4.2) which is asymmetrical along the  $z$  axis. The same number of prints was performed with Honeycomb and Rectilinear fill audio sliced to 90s and 120s each and CTh set to 35. The ROC curve of these experiments are shown in Figure 4.7. Each sample consists of nine target prints and sixty malicious prints, and the same data is used for the 90 second audio slice length as the 120 second slice.

Upon investigation of the G-code, the only repeated lines were those that define the nozzle speed at the beginning and do not include extrusion. Furthermore, there are no blocks of G-code or layers that are entirely repeated verbatim. This is suspected to contribute greatly to the increased performance seen in Figure 4.7. Here, least accuracy is 98.52% which is suitable for verification purposes. Between the 120 second and 90 second slice lengths, we see little change in performance. Although audio classification is shown here to be effective in identifying malicious prints, it is still susceptible to a non-negligible false positive rate.

By introducing data from the spatial layer, false positives may be reduced. For instance,

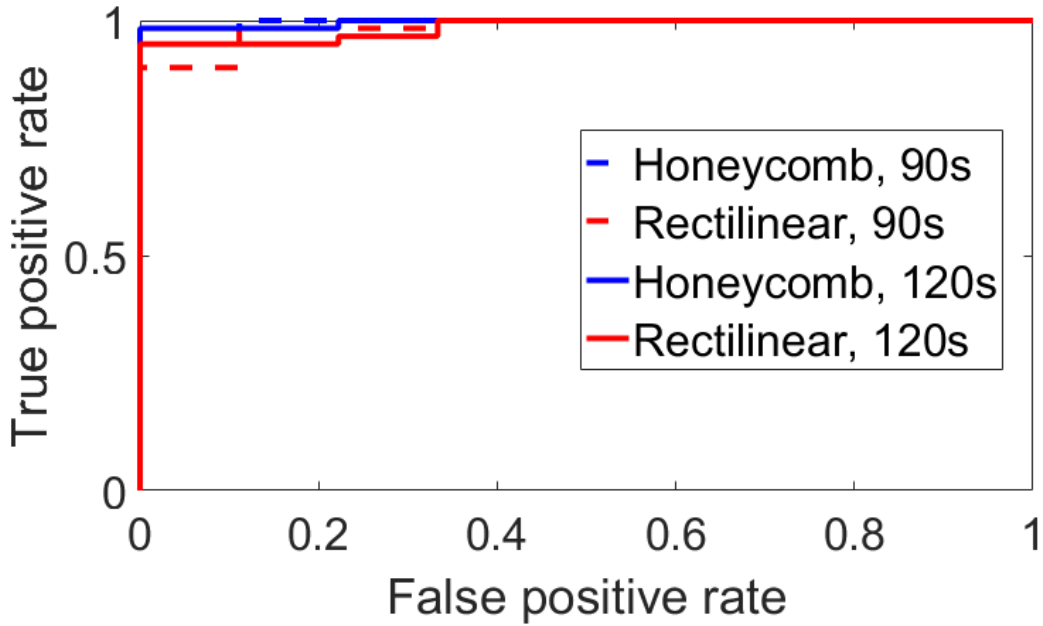


Figure 4.7: ROC curves for Top Hat.

Figure 4.8 compares the frequency domain data from the  $x$ ,  $y$ , and  $z$  axes of the 40% Honeycomb and 40% Rectilinear fills from Figure 4.5. Here, we see a significant difference between the two prints. Each frequency response has a similar shape, but the major peaks of the 40% Rectilinear fill are shifted to the right because the back-and-forth motion is not impeded by the creation of small Honeycomb structures.

For classification, the four most prominent peaks are used as features along with their locations. We conducted a test in which the benign print was chosen to be the 20% Rectilinear disk shown above. All other prints were considered malicious. With this, we had 10 benign prints and 12 malicious prints. Training using the linear regression model, an AUROC of 1 was achieved in differentiating between malicious and benign prints.

While the spatial sensing layer is primarily for the purpose of print visualization, its role in conjunction with the acoustic layer allows for high accuracy in detecting malicious prints.



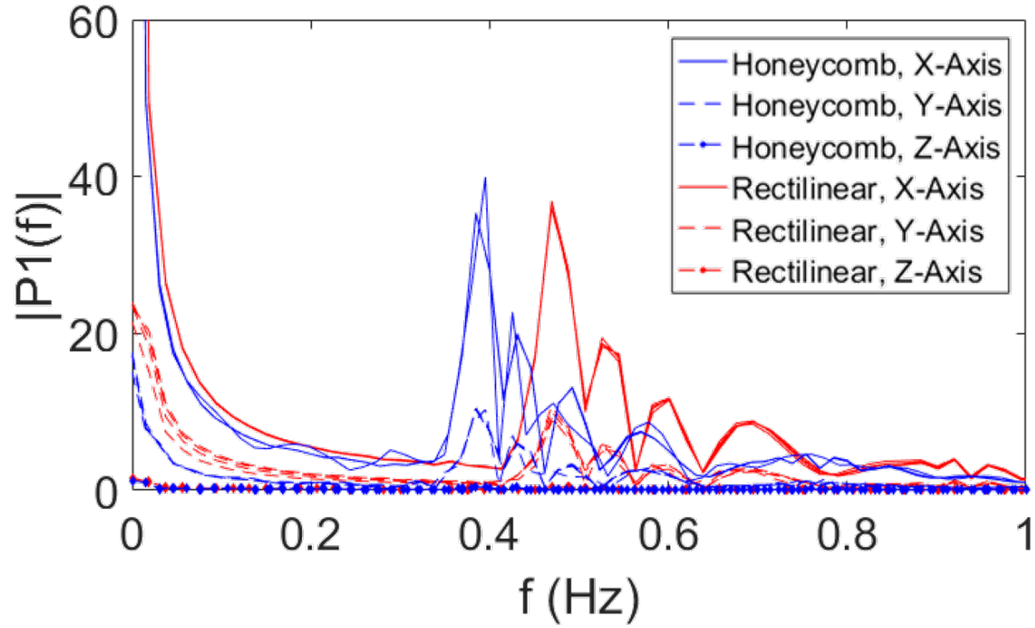


Figure 4.8: Comparison of the frequency response between a single layer of Honeycomb 40% fill and Rectilinear 40% fill.

#### 4.2.2 Varied Printer Models

In order to understand the effectiveness of audio classification for print verification on different printer models, several prints were performed on a Lulzbot TazMini and Orion Delta. Acoustic data recordings are obtained using the same microphone. In each print, a Top Hat design was printed and the audio was sliced to 120s. The optimized CTh for the TazMini, Orion Delta, and Taz6 are 150, 20, and 35 respectively. The CTh values differ between printers because each has different components that produce varying levels of noise as they print. The ROC curve results are shown in Figure 4.9. Because the Honeycomb and Rectilinear fill patterns are considered together, each data set consists of 18 target prints and 120 malicious prints. Because the classification accuracy does not fall an AUROC of 0.9542 in these tests, the acoustic verification method is generalizable, given prior optimization of CTh.

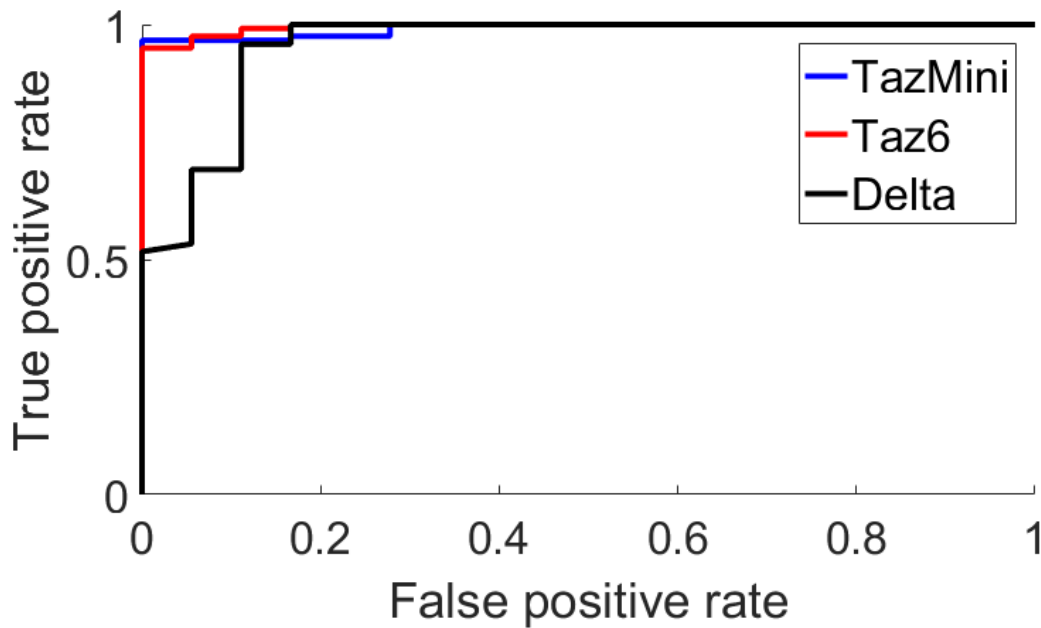


Figure 4.9: ROC curves for top hat design printed using a TazMini, Orion Delta, and Taz6 print. Print audio sliced to 120 seconds and the confidence threshold is 150, 20, and 35 respectively.

#### 4.2.3 Detecting Extrusion

One likely attack on an AM system would be to have the system appear to be running while not extruding material. The result would be a waste of time and money. To determine if the acoustics of the system could detect such an attack, the Top Hat design with a 40% Honeycomb fill was printed 12 times with extrusion and 12 times without. The training data was a print of the former with the audio split to 120 seconds. CTh is set to zero, so the scores are calculated based on indexing alone. The printer used here was a Rostock Max v3. The same recording equipment was used, but this time the microphone was directed at the motors rather than the nozzle.

The results are shown in Figure 4.10. The top plot indicates average scores calculated for a print with extrusion and the bottom without. A clear difference of two orders of magnitude can be seen between scores with and without extrusion within the first 120 seconds.

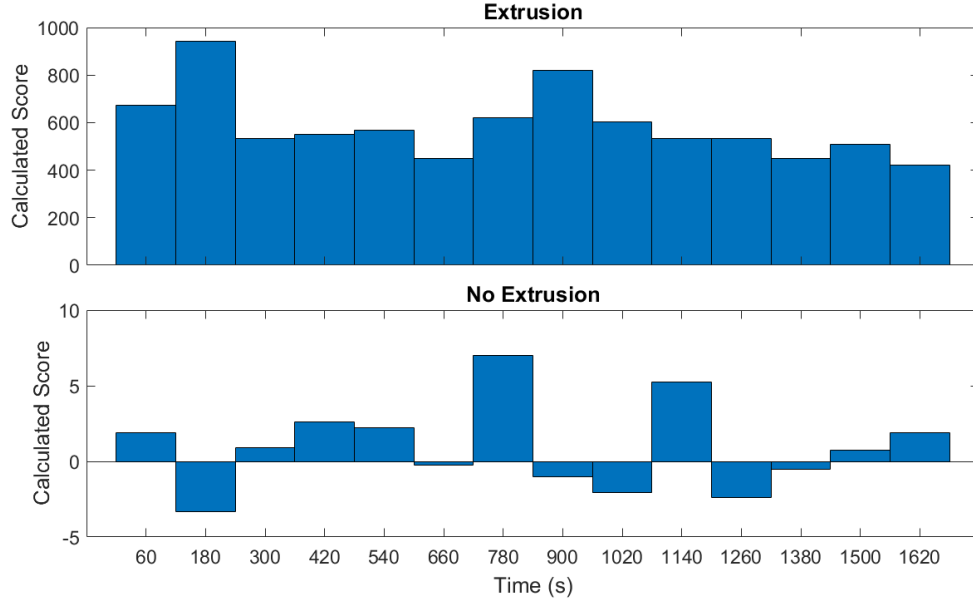


Figure 4.10: Calculated scores differentiating prints that extrude material and those that do not.

#### 4.2.4 Classification with Minimal Change

Another experiment was performed using the Rostock Max v3 in which the steps in density changes for the prints was reduced from 20% to 10%. The fill designs were Honeycomb and Rectilinear, and the densities were 20%, 30%, and 40% each. 12 copies of each print were performed with the extruder off so that prints could be automated in succession. Overall, this resulted in a data set of 72 benign prints and 360 malicious prints where any benign print can be considered a malicious print for another training set. The Cth was set equal to zero.

The results are shown in Figure 4.11. We see that even small changes that are unlikely to cause major structural change can be detected through audio classification with an AUROC value of 0.9418.

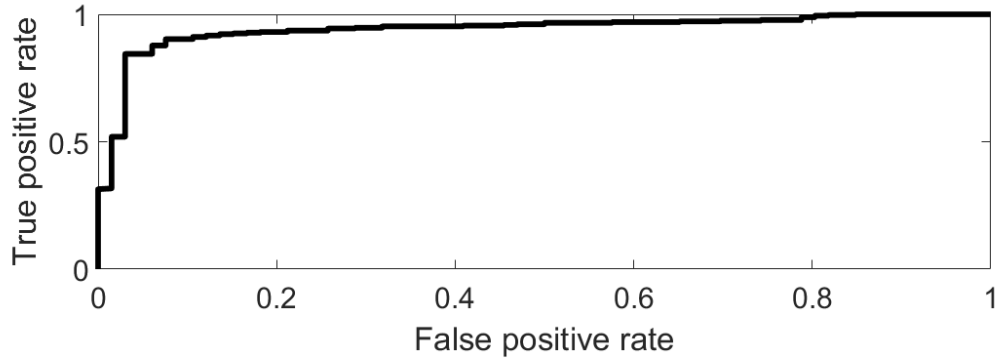


Figure 4.11: ROC curve for malicious print detection between Honeycomb and Rectilinear prints with 20%, 30% and 40% fill density.

#### 4.2.5 Classification in Noisy Environments

Other experiments were conducted using an Afina H40 3D Printer with an eBoTrade Digital Voice Recorder wide-range microphone. This setup was in a noisy university makerspace with people talking near the printer. In this experiment, the classification accuracy suffered greatly (near AUROC = .5). Because it is shown that acoustic verification is useful on different types of printers above, we assume that the loss of classification accuracy is due to the noise in the environment. Also, because the microphone was wide range and not directional, the talking near the printer can be clearly heard. Therefore, in the implementation of this verification scheme it is important to use a directional microphone and noise isolation as much as possible.

#### 4.2.6 Visualization of Malicious Prints

When a potentially malicious print is identified as described above, it is important to have the capability to visualize the potential threat. This visualization must be independent of the intended G-code which may be interpreted differently by malicious firmware.

**Real-Time Visualization.** In the event that a potential malicious print is identified, a user has the capability of viewing it in progress through spatial sensing as seen in Figure 4.5. By viewing the layer in progress, significant fill pattern changes such as those between

the 20% Honeycomb and 20% Rectilinear fill are obvious. However, less obvious changes made to the print such as those between the 40% Honeycomb and Rectilinear fills are identifiable through FFT Analysis as in Figure 4.8. This is particularly true, as will be shown in subsection 4.2.7, if the user has access to the frequency response of a reference print.

While the spatial sensing layer is useful for identifying the type of fill pattern that is being maliciously generated, it is less useful for identifying if the design itself has been altered due to the warping that occurs in the data. This, however, is an easy issue to solve through the use of a web cam which can easily identify the shape of the design. In this sense, it may seem that spatial sensing may be replaced altogether by a web cam, but it is important that the web cam uses far more data and does not readily provide information about the frequency response.

#### 4.2.7 Case Study: Prosthetic Knee

A model of the tibial component of a prosthetic knee implant was used as a design for a use case test. Prosthetics differ slightly between patients, so we assume that malicious print identification is performed periodically with a known standard prosthetic design. In this scenario, real-time visualization may still be performed on each print.

**Error Identification.** The acoustic verification results are shown in Figure 4.12 which shows the confidence values of both the target print and the malicious print. These results are gathered using the same technique as those described in section 4.1 with audio slices of length 120s and  $C_{Th} = 0$ . By setting  $C_{Th} = 0$ , we see that a positive error classification can be made within the first 360s of the print or the first 4% of the total known print time by only observing out-of-sequence index classifications. The  $C_{Th}$  may be set to anything less than 18 without causing a false positive. Here, acoustic error detection itself saves over 2 hours of print time and prevents a potentially harmful print from being completed. A detailed table of the results shown here can be found in appendix section 6.2.

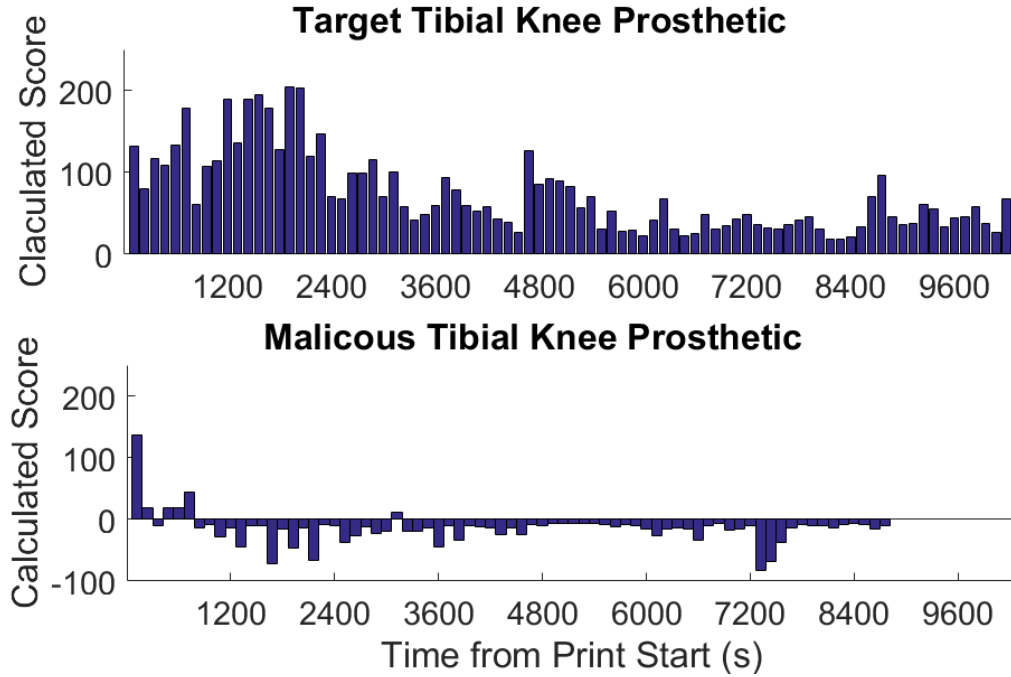


Figure 4.12: Comparison of acoustic print classification for target tibial prosthetic with 60% Rectilinear Fill (Top) vs. malicious 20% Honeycomb Fill (bottom). CTh = 0.

In Figure 4.13, the FFT of a benign print and a malicious print are compared to a training print. Similar to Figure 4.8, the malicious print shows a different frequency response near 0.2Hz as highlighted by the lower box. The upper box highlights the closeness of the peaks between the training and benign prints and the difference between those and the malicious print. The full print of the object requires 111 layers, so it would take less 1% of the time of the total print to identify the erroneous pattern once it begins.

**Real-Time Visualization.** In this test, the target print uses a 60% Rectilinear fill and the malicious print uses a 20% Honeycomb fill. In the attack, the visualization of the intended G-code remains unaltered for the user while the instructions sent to the printer are altered. The consequences of this attack would be to cause accelerated wear in the implant causing pain and financial loss for the victim who has the implant.

Due to the availability of the experimental setup, a single layer of each print was performed by the Dobot Magician for the visualization tests. The same designs and G-code

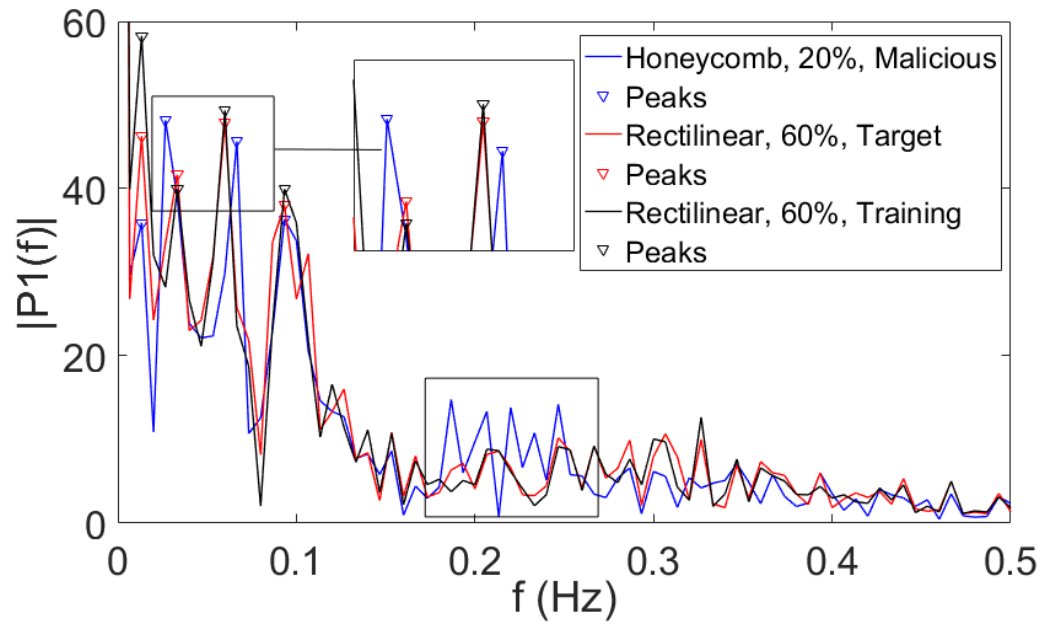


Figure 4.13: Comparison of x-axis frequency response for a layer of the tibial knee implant design.

was used as in the previous acoustic classification tests. It should also be noted that both acoustic and spatial verification would ideally be performed in tandem, but for testing purposes here, they are not. As seen in Figure 4.4, the linear potentiometer is secured to a stationary table. However, it is feasible for it to be attached to a mobile base to measure the relative motion between the base and the nozzle.

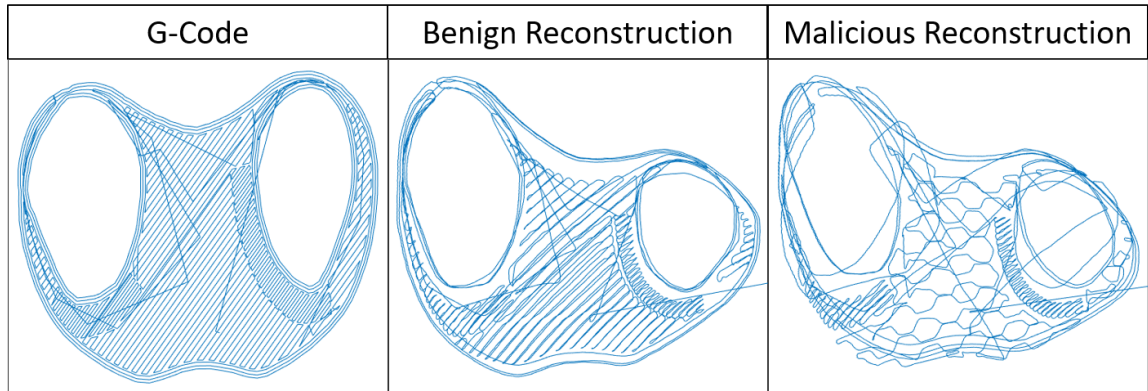


Figure 4.14: Comparison of benign and malicious tibial knee implant prints.

Figure 4.14 shows the spatial verification visualization of, in order of left to right, a

G-code visualization of the training print, a spatial reconstruction of the benign print, and a spatial reconstruction of the malicious print. It is clear that the recreated benign print uses a rectilinear fill at approximately the correct density while the malicious print differs significantly from the intended G-code. Due to the warping that occurs in the spatial reconstruction, a user would not be made aware if the shape of the print were altered by using this method alone.

### **4.3 Implementation**

The software that controls the microphone and transmission of acoustic data will ultimately be integrated into the printer firmware. Because chapter 3 describes the target of the attack as the firmware itself, it is important to ensure that it would be infeasible to replay or spoof the data used for audio classification.

The first way to do this would be to ensure that the actual classification takes place at the user end. If the classification takes place on the firmware and a simple response is sent to the user indicating a correct or incorrect classification, the system is susceptible to spoofing. If the raw acoustic data is sent, the attacker would have to hard-code a replay of the benign audio or somehow generate the audio based on the G-Code itself. However, this creates a new challenge of increased network traffic. If print audio is sampled at 44100Hz at a 16 bit word length and the print lasts 60 minutes, approximately 310MB of additional data must be sent to the user. The use of compressed audio is beyond the scope of this thesis, but future work will explore its usefulness in potentially decreasing the load on the network.

A theorized attack on this system may include a type of replay attack where a recording of a benign print is played on a speaker during a malicious print to fool the classifier. To test this, a recording of a Top Hat design with the Honeycomb fill pattern at 40% density was played on a JBL Charge 2 speaker near the microphone on the Rostock Max v3 printer. The microphone was not moved from the location where the original recording was performed.



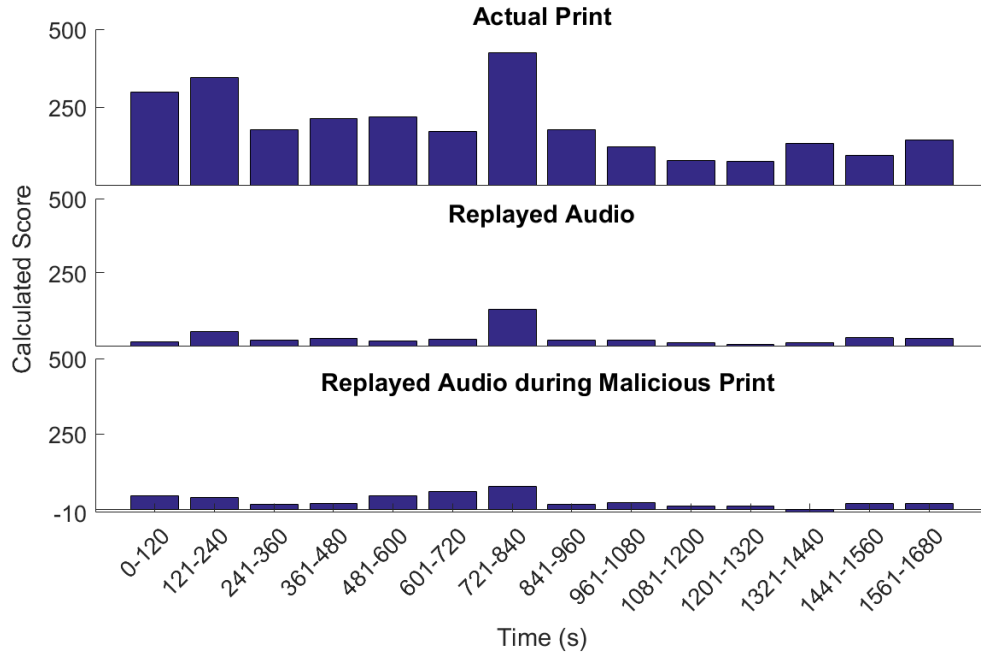


Figure 4.15: Classification data from attempted physical replay attack.

Two tests were conducted. In the first, the recording of the benign print was played with the 3D printer off. In the second, the benign print audio was played on the speaker while the 3D printer printed a Top Hat design with the Rectilinear fill pattern at 20% as the malicious design. Figure 4.15 shows the results. The top plot are the classification results from when the benign print was actually printed by the printer. The middle plot is the result when the audio was replayed with the printer off. The bottom plot is the results when both the replay attack and malicious print took place in tandem. In each attack scenario, the index values of the classifications were consistent with the training set with the exception of 1321-1440 seconds into the print in the third experiment. However, the confidence values were significantly lower in the attack scenarios than when the true benign print is classified. This is likely due to environmental noise or changes to the acoustic waveforms when the audio is played through the speaker. The data presented here contains only one trial, so future work will have to explore the possibility of physical replay attacks further.

## CHAPTER 5

### ELECTRICAL GRID INTRUSION DETECTION

#### 5.1 Switching Detection Methods

The experiments performed in this thesis focus on distribution substations on the electrical grid. Substations contain transformers that change voltage levels from high for transmission to low for distribution or vice versa. Figure 5.1 illustrates the generation, transmission, distribution system with the substation highlighted by the box. Substations also contain monitoring equipment, and are the point at which electrical grid circuits may be opened and closed. This is also referred to as switching.

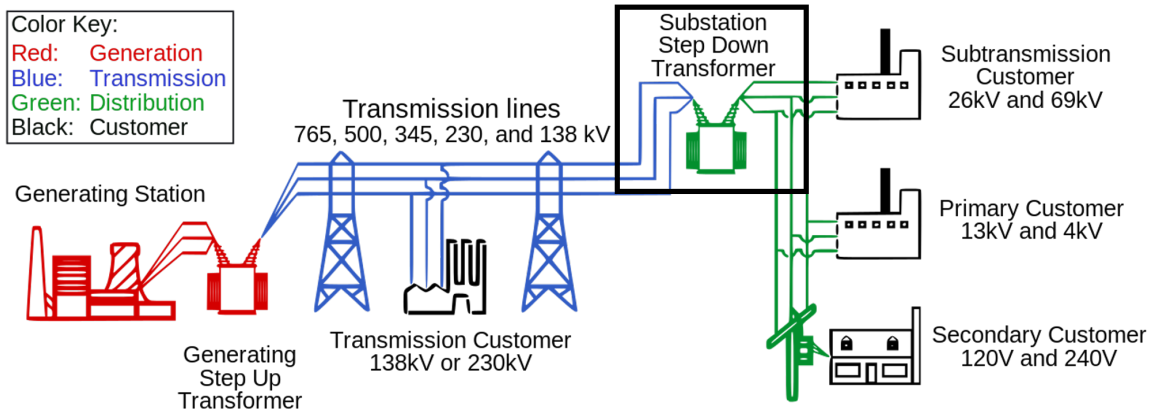


Figure 5.1: Diagram of a simple electrical grid [31]

In order to perform maintenance or to upgrade equipment, substations are often taken off-line temporarily by power companies. To do this, switching is performed such that the circuits that are connected to the substation are open. Before the switching, parallel circuits are established so that no outages are experienced by the clients that are connected to the distribution lines.

To measure power system activity, the AWESOME receiver was connected to a bundled solenoid antenna. The antenna consisted of 30.18m of 20 AWG copper wire wrapped

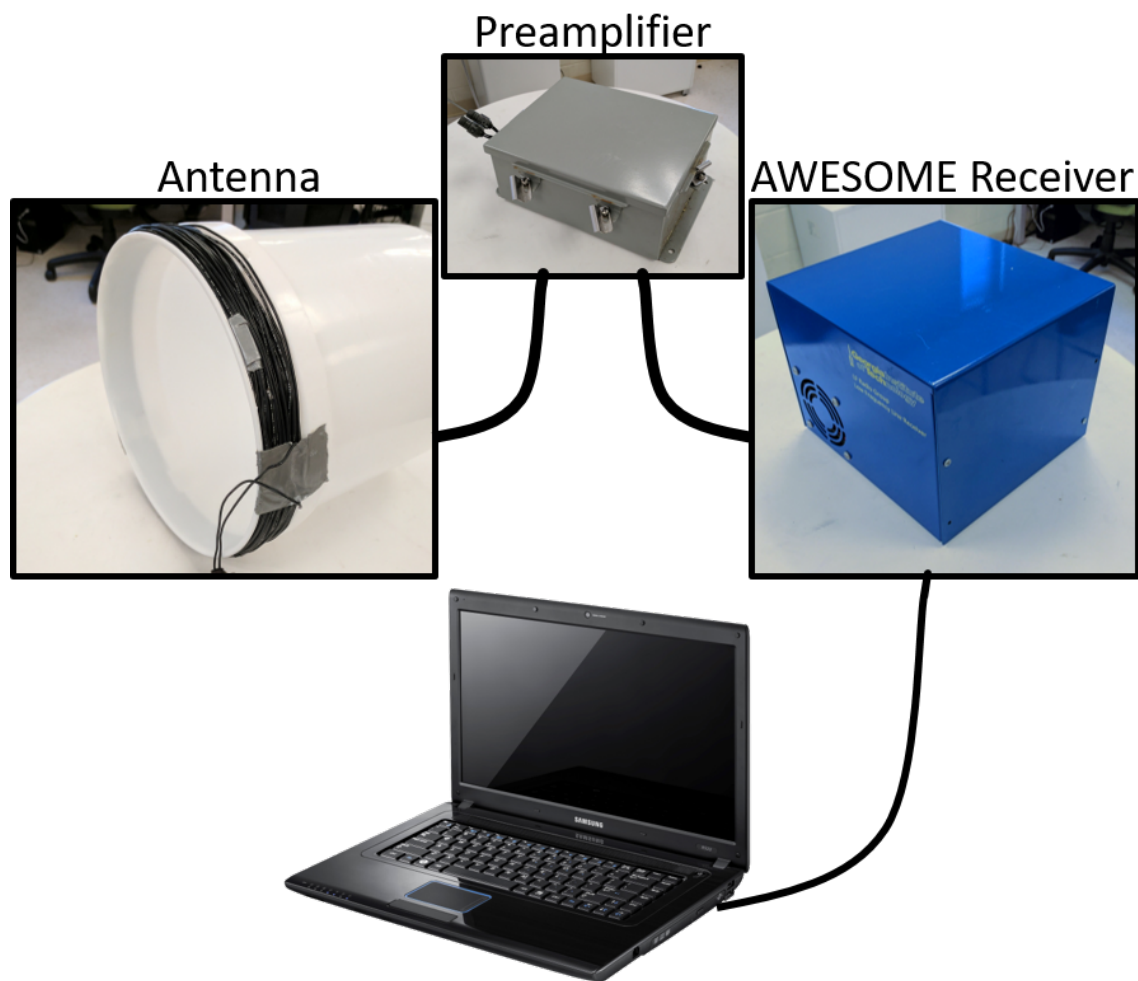


Figure 5.2: AWESOME receiver and antenna setup.

around a 23cm diameter circle as shown in Figure 5.2. The signal from the antenna is amplified by the pre-amplifier and then sent to the AWESOME for processing. The result is a signal sampled at 1MHz measuring the magnetic field.

The 60Hz background frequency caused by the normal operation of the power system is then filtered out along with any man-made VLF transmission signals [25]. Man-made VLF transmission signals are used for specialized communication systems and their frequencies at which they operate are also subtracted from the data [32]. A sample of resulting data is shown in Figure 5.3. This data was collected at 3854°2.65'N, 7548°44.06'W on March

28, 2017. The red arrows indicate precise timestamps of lightning detected by NLDN. Because of the high time sensitivity of both NLDN and AWESOME, propagation delay of the lightning event must be considered. The greatest distance between any two locations in the 48 contiguous United States is 4,669 km [33]. Sferics are known to travel  $> .99c$ , so any signal caused by a lightning event will fall within approximately  $15,600 \pm 0.5$  microseconds of the event as detected by NLDN. In Figure 5.3, we see two groups of lightning events. We can clearly see that the prominent peaks align with the lightning events. The series of events that takes place between 10:52:29-30 is located 76km from the receiver and the event between 10:52:33-34 is 53.9 km away. The lower plot shows the signal of the lightning event highlighted by the blue box in the upper plot.

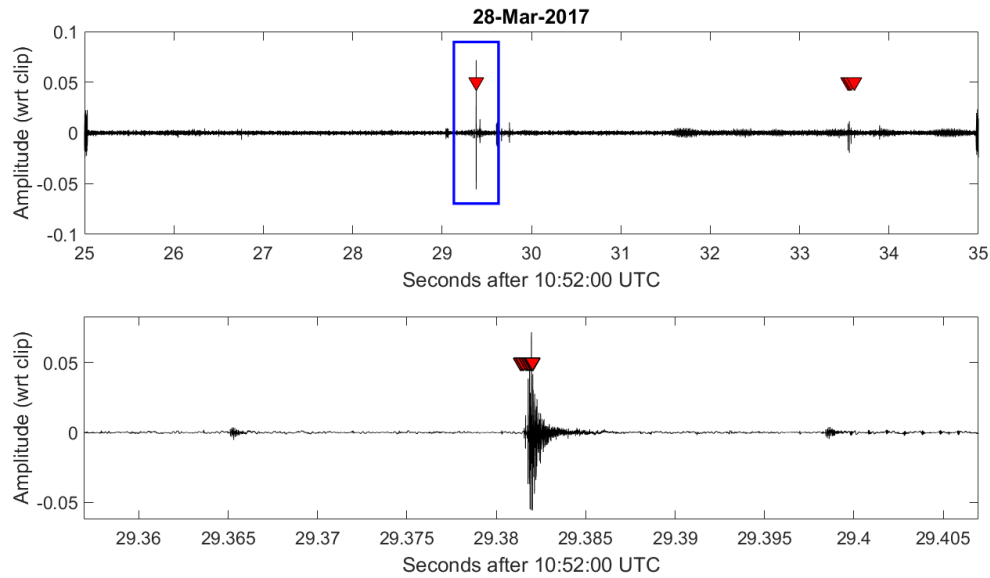


Figure 5.3: Sample recorded LF data with two groups of lightning strikes.

Measurements were taken in collaboration with a utility company during various switching events that were necessary for maintenance and repairs. The antenna was positioned inside the substation fence and such that the face was perpendicular to the outgoing current and approximately ten feet away. The power for the receiver equipment came from a car battery connected to an alternator, and no physical connection was made to any substation

equipment. The observed events included the opening of four circuits (two circuits at two separate substations) and the closing of one of the previous circuits (the other three would remain open beyond the observation time).

## 5.2 Evaluation

The switching events that were recorded for this experiment took place at two substations, referred to as Sub1 and Sub2, which are located at 38°9'8.46"N, 75°41'34.82"W and 38°54'2.65"N, 75°48'44.06"W respectively. Two circuits were opened at each substation. The circuits here are referred to as C1, C2, C3, and C4, where C1 and C2 are connected to Sub1, and C3 and C4 are connected to Sub2.. The antenna for the AWESOME receiver was placed inside the substation boundaries with the open face perpendicular to the outgoing distribution lines.

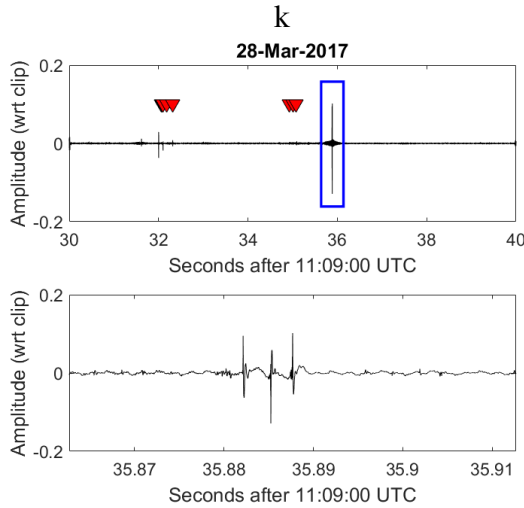


Figure 5.4: Current C1 opened.

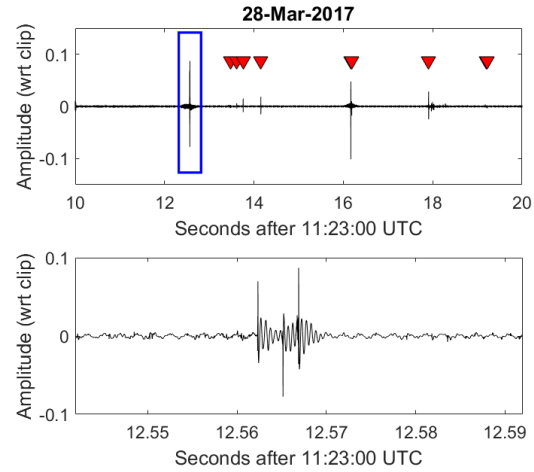


Figure 5.5: Current C2 opened.

The SCADA records show that C1 and C2 opened at 11:09:39 UTC and 11:23:18 UTC respectively on March 28, 2017. Figure 5.4 and Figure 5.5 show the time series data for these events. The top plot shows the time series magnetic field signal with the amplitude scaled to a fraction of the value where the data would clip. NLDN lightning events are indicated by red triangles. The switching events are suspected to occur at 11:09:35.88

UTC and 11:23:12.56 UTC as indicated by the boxed feature. These are seconds before the events are recorded by SCADA. At these times, the magnetic field signal reaches a maximum when no lightning events are recorded by NLDN. The bottom plot shows a detail of the feature. Here, we see three distinct peaks which is consistent with the presence of 3 phases in the power line. This signal is significantly different from previously studied typical lightning signatures and may, in principle, be easily distinguished [34].

C1 is closed again and the event is recorded by SCADA at 11:47:44 UTC. Figure 5.6 shows the data during this time and highlights the suspected associated signal at 11:47:44.02 UTC as indicated by the box. The bottom plot which shows the feature in detail does not show the three clear peaks. This may be due to the three-phase power requiring some amount of time to stabilize after the circuit is closed.

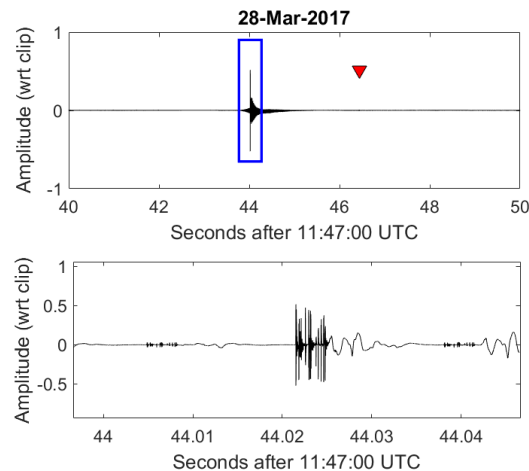


Figure 5.6: Current C1 closed.

At Sub2, we see similar results when circuits C3 and C4 are opened. The SCADA data shows these circuits opening at 10:17:38 UTC and 10:18:58 UTC respectively on March 27, 2017. As shown in Figure 5.7 and Figure 5.8, features likely to correspond to the actual events occur at 10:17:36.28 UTC and 10:18:53.08 UTC respectively. The C4 opening event shows the same three clear peaks as C1 and C2 did. However, the C3 opening event is surrounded by several recorded lightning strikes which cause sferics that distort the data. Still, the highest peak in the ten second window is not accompanied directly

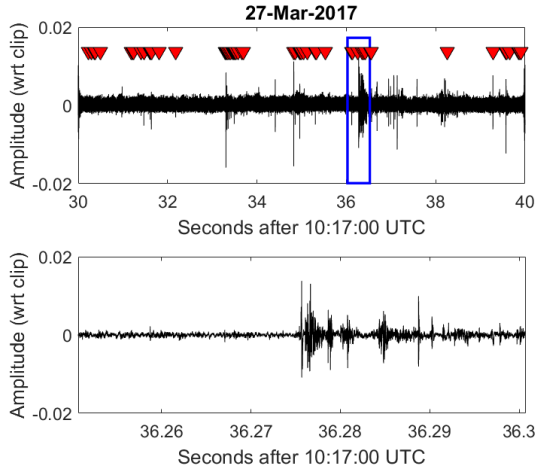


Figure 5.7: Current C3 opened.

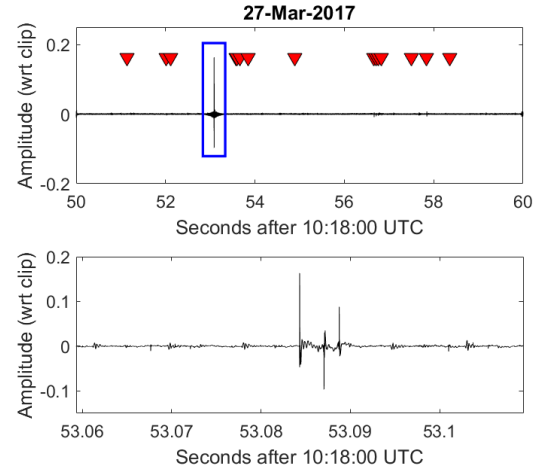


Figure 5.8: Current C4 opened.

by any individual NLDN lightning events. The lightning strike closest to the time of the peak occurs at 10:17:36.25 UTC at a location 1,280km from the site of the antenna. In order for the signal seen to have been caused by the lightning strike, the signal would have to be going significantly slower than the known speed of sferic propagation [24, 25].

### 5.2.1 Network Traffic Comparison

In order to use low frequency switching detection as a form of intrusion detection, it is important that the detected signals correspond to real network traffic. If a switch is detected with the receiver, but no corresponding instructions are sent, or vice versa, it could indicate an intrusion that has managed to bypass network-based intrusion detection. This knowledge can allow for an early start to the recovery process from the attack.

Network traffic was captured during the switching of the C1 and C2 circuits and is shown in Figure 5.9. Analysis of the DNP 3.0 protocol shows separate 'DO' commands at both 11:09:35 and 11:23:12. These commands were confirmed by the power company that owns the network to open all three breakers for the C1 and C2 circuits respectively. Furthermore, the packets captured during 11:09:39 and 11:23:18 show the communication with the SCADA to log the events.

C1 Open					
11:09:35.387	██████.20.22	██████.0.11	DNP 3.0	89 from 1024 to 48, Select	
	██████.0.11	██████.20.22	TCP	60 20000 → 65528 [ACK] Seq=10996 Ack=3925 Win=16384 Len=0	
11:09:35.422	██████.0.11	██████.20.22	DNP 3.0	91 from 48 to 1024, Response	
11:09:35.437	██████.20.22	██████.0.11	DNP 3.0	89 from 1024 to 48, Operate	
11:09:35.455	██████.0.11	██████.20.22	DNP 3.0	91 from 48 to 1024, Response	
11:09:39.558	██████.20.22	██████.0.11	DNP 3.0	78 from 1024 to 48, Read, Class 123	
	██████.0.11	██████.20.22	TCP	60 20000 → 65528 [ACK] Seq=11070 Ack=3984 Win=16384 Len=0	
11:09:39.570	██████.0.11	██████.20.22	DNP 3.0	274 from 48 to 1024, Response	
11:09:39.609	██████.20.22	██████.0.11	DNP 3.0	69 from 1024 to 48, Confirm	

C2 Open					
11:23:12.307	██████.20.22	██████.0.11	DNP 3.0	89 from 1024 to 48, Select	
	██████.0.11	██████.20.22	TCP	60 20000 → 65528 [ACK] Seq=6007 Ack=2679 Win=16384 Len=0	
11:23:12.310	██████.0.11	██████.20.22	DNP 3.0	91 from 48 to 1024, Response	
11:23:12.356	██████.20.22	██████.0.11	DNP 3.0	89 from 1024 to 48, Operate	
11:23:12.377	██████.0.11	██████.20.22	DNP 3.0	91 from 48 to 1024, Response	
11:23:18.567	██████.20.22	██████.0.11	DNP 3.0	78 from 1024 to 48, Read, Class 123	
	██████.0.11	██████.20.22	TCP	60 20000 → 65528 [ACK] Seq=6098 Ack=2762 Win=16384 Len=0	
11:23:18.600	██████.0.11	██████.20.22	DNP 3.0	286 from 48 to 1024, Response	
11:23:18.617	██████.20.22	██████.0.11	DNP 3.0	69 from 1024 to 48, Confirm	

Figure 5.9: Network traffic for currents C1 and C2 opened.

## 5.2.2 Electrical Grid Event Detection at a Distance

As part of the AWESOME project, several receivers are installed throughout the United States and around the world [35]. These receivers are set up with much larger antennas and are able to detect sferics that propagate from around the earth. Finally, these receivers are each connected to two antennas - one oriented to detect signals propagating in the North-South direction, and one for East-West.

One such receiver was located at 3916'42.60"N, 7534'52.68"W, approximately 125km North of the Sub2 location as shown in Figure 5.10. After the circuits C3 and C4 were opened, the transformer on the circuit was discharged. This caused a large visible arc that disrupted the on-site receiver signal at 10:21:23 as shown in the top plot of Figure 5.11. We investigated the signal from the aforementioned AWESOME network receiver which is shown in the bottom plot. Just after 10:21:23, a large peak appears in the North-South portion of the signal, but does not appear in the East-West signal.

One lightning event is identified just before 10:21:23 UTC which is 1,280km away from Sub2. Figure 5.12 shows a detail of the external receiver around 10:21:23. The white triangle indicates the time of the distant lightning, and the black triangle indicates where the



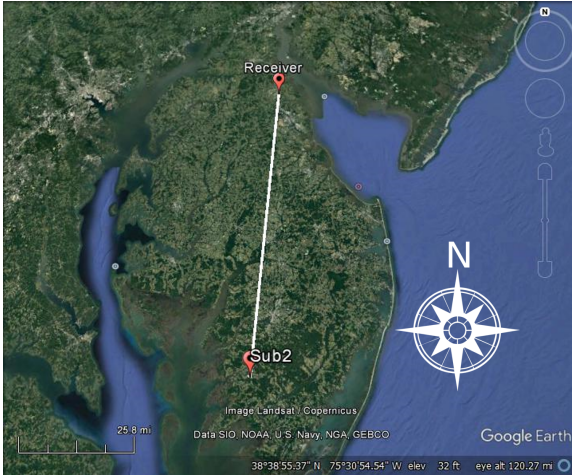


Figure 5.10: Distance from Sub2 to AWE-SOME network receiver.

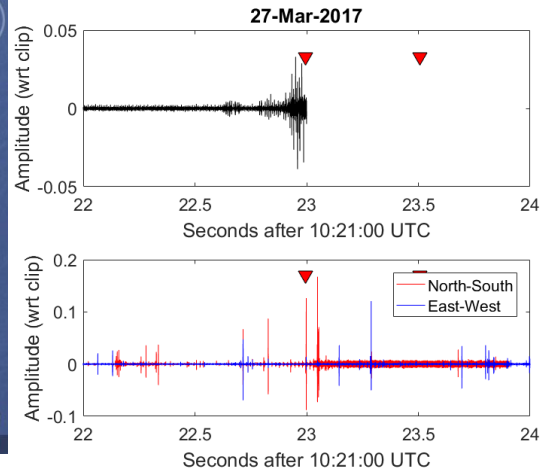


Figure 5.11: Transformer switch.

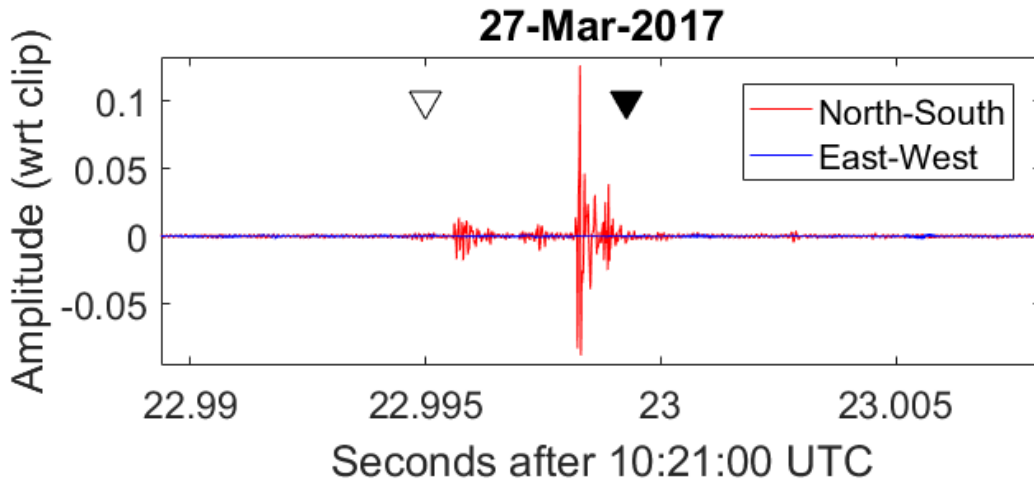


Figure 5.12: Transformer switch detail.

signal would be detected assuming that the signal travels at the speed of light in a vacuum. Here we see that in order for the suspect signal to be caused by the lightning, the signal would have to be moving at a approximately 130% the speed of light. This, along with the fact that the North-South portion of the signal reacts much more strongly than the East-West portion, indicates that the peak may be a result of the arc caused by the transformer switching.

On April 12, 2016 an incident occurred in which a squirrel caused a short circuit which resulted in a major outage in a third substation (Sub3) located 64.6 km Southwest of the

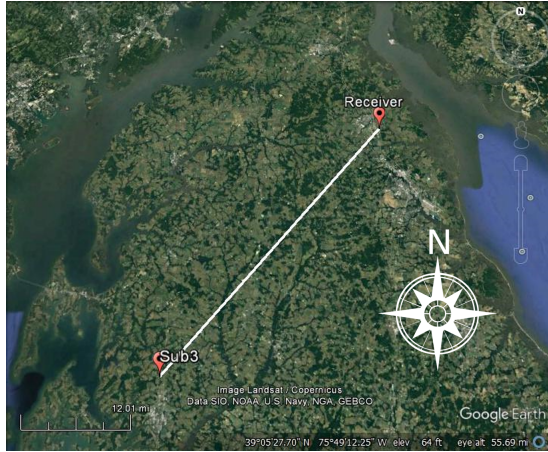


Figure 5.13: Distance from Sub3 to AWE-outage. SOME network receiver.

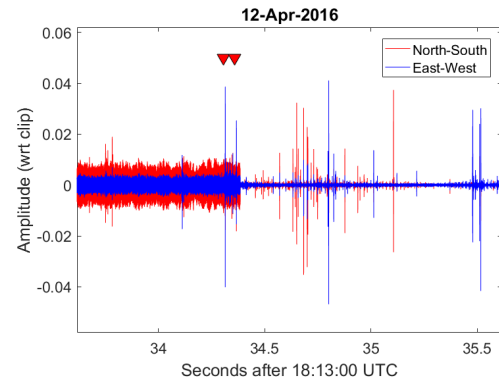


Figure 5.14: AWESOME data during Sub3

AWESOME network receiver. The path is shown in Figure 5.13. The incident caused a large visible arc and was recorded by SCADA at 18:13:35. The corresponding data is shown in Figure 5.14. Here, we see a significant change in the average signal amplitude just before 18:13:34.5. This effect is caused by the removal of the 60Hz signal and harmonics which dominates the RMS floor noise.

The detection of power grid activity with distant receivers is plausible with the use of waveform classification to distinguish grid events from natural events. Switching events are shown to appear significantly different from typical lightning events, and may be identified amidst the multitude of signals detected by the large antennas. Furthermore, it is possible that arcs caused by outages or switching procedures may be detectable as well. This is due to the fact that in neither case above was the event mistakenly identified by NLDN as a lightning strike despite the similar mechanics between lightning and arcing.

### 5.3 Implementation

Using waveform classification and modeling, it is feasible for an array of AWESOME receivers to be positioned to monitor electrical grid activity by differentiating signals caused by the grid from those caused by natural events such as lightning.

The primary implementation challenge for the power system intrusion detection system is again the threat of a replay attack. While spoofing network traffic, an attacker may be able to replay benign data on the network as if it were from the AWESOME receiver. The main defense against this would be to use the randomness of natural atmospheric signals to our advantage.

For instance, if a set of data is being replayed by an attacker that contains no lightning events, a nearby lightning event could occur that is recorded by NLDN with no corresponding signal in the receiver data. This could alert the user to malicious activity on the network. Additionally, if multiple receivers are positioned at nearby substations, any natural atmospheric signal in the nearby area would be expected to be represented similarly in the data from each receiver. If an attacker were to remain undetected, a replay attack would have to be performed in tandem for multiple receivers during very fair weather.

A challenge in this intrusion detection method is again the increased network traffic. The data collected here is sampled at 1MHz with a 16 bit word length. This would equate to over 7GB per hour per channel of data. However, this traffic may be reduced by on-site processing which only transmits data for certain detected events.

## **CHAPTER 6**

### **CONCLUSION AND FUTURE WORK**

#### **6.1 Conclusion**

In this thesis, we show two methods of physical signal-based intrusion detection.

For the sector of Additive Manufacturing, audio classification and spatial position analysis is used to verify that the internal or external structure of a print has not been altered by malicious firmware. We show that a change in the fill density, fill structure, or the presence or lack of extrusion can cause a misclassification of an intended print and thus alert the user to intrusion.

For the electrical grid sector, we show that analysis of LF magnetic field signals can identify switching events. We show how these signals can be differentiated from natural events such as lightning and how distant receivers may be able to detect significant events such as arcing.

Finally, we discuss how these methods may be implemented in a practical way and the challenges that may be faced in their implementation. A primary challenge is found to be the large amount of data that must be sent over the network in order to process it on the user end.

#### **6.2 Future Work**

Future work for audio classification intrusion detection will focus on its integration with 3D printer firmware. Ultimately, the microphone and its controller should work in conjunction with the mechanics of the 3D printer and be able to communicate with the user all at once. Also, the method will be expanded to different types of 3D printers such as selective laser sintering which is used for printing metal objects. Also, the method can be expanded to

other processes which have acoustic signals as a significant physical byproduct.

Future work in spatial sensing for AM will focus on the ability to directly compare G-Code instructions to spatial reconstruction with on-board sensors. Many 3D printers are equipped with accelerometers for calibration purposes. These may be good candidates for providing useful spatial data.

Future work in LF sensing on the electrical grid will be further integrated into currently established procedures. Real-time signal processing will be developed in order to identify grid activity without the need for large amounts of data being sent over the network. Finally, signal analysis will be performed to see if LF radio signals from the electrical grid can convey information about changes in current and power.

# **Appendices**

# APPENDIX A

## DETAILED RESULTS OF ACOUSTIC CLASSIFICATION ON TIBIAL KNEE

### PROSTHETIC

Tibial Knee Prosthetic Classification, Trained with Rectilinear Fill, 60% Density															
60% Rectilinear Fill				20% Honeycomb Fill				60% Rectilinear Fill				20% Honeycomb Fill			
Index Value	Classification Result	Confidence	Index Value	Classification Result	Confidence	Index Value	Classification Result	Confidence	Index Value	Classification Result	Confidence				
0	Taz6Tbia Rectilinear 60 T(10)	132	Taz6Tbia Rectilinear 60 T(10)	137	43	Taz6Tbia Rectilinear 60 T(43)	57	Taz6Tbia Rectilinear 60 T(53)	7	Taz6Tbia Rectilinear 60 T(53)	7				
1	Taz6Tbia Rectilinear 60 T(11)	80	Taz6Tbia Rectilinear 60 T(11)	19	44	Taz6Tbia Rectilinear 60 T(44)	70	Taz6Tbia Rectilinear 60 T(55)	8	Taz6Tbia Rectilinear 60 T(55)	8				
2	Taz6Tbia Rectilinear 60 T(12)	117	Taz6Tbia Rectilinear 60 T(12)	10	45	Taz6Tbia Rectilinear 60 T(45)	31	Taz6Tbia Rectilinear 60 T(56)	12	Taz6Tbia Rectilinear 60 T(56)	12				
3	Taz6Tbia Rectilinear 60 T(13)	108	Taz6Tbia Rectilinear 60 T(13)	19	46	Taz6Tbia Rectilinear 60 T(46)	53	Taz6Tbia Rectilinear 60 T(57)	9	Taz6Tbia Rectilinear 60 T(57)	9				
4	Taz6Tbia Rectilinear 60 T(14)	133	Taz6Tbia Rectilinear 60 T(14)	18	47	Taz6Tbia Rectilinear 60 T(47)	28	Taz6Tbia Rectilinear 60 T(58)	10	Taz6Tbia Rectilinear 60 T(58)	10				
5	Taz6Tbia Rectilinear 60 T(15)	178	Taz6Tbia Rectilinear 60 T(15)	45	48	Taz6Tbia Rectilinear 60 T(48)	29	Taz6Tbia Rectilinear 60 T(59)	15	Taz6Tbia Rectilinear 60 T(59)	15				
6	Taz6Tbia Rectilinear 60 T(16)	61	Taz6Tbia Rectilinear 60 T(16)	13	49	Taz6Tbia Rectilinear 60 T(49)	23	Taz6Tbia Rectilinear 60 T(60)	27	Taz6Tbia Rectilinear 60 T(60)	27				
7	Taz6Tbia Rectilinear 60 T(17)	107	Taz6Tbia Rectilinear 60 T(17)	9	50	Taz6Tbia Rectilinear 60 T(50)	41	Taz6Tbia Rectilinear 60 T(61)	15	Taz6Tbia Rectilinear 60 T(61)	15				
8	Taz6Tbia Rectilinear 60 T(18)	114	Taz6Tbia Rectilinear 60 T(18)	28	51	Taz6Tbia Rectilinear 60 T(51)	67	Taz6Tbia Rectilinear 60 T(62)	14	Taz6Tbia Rectilinear 60 T(62)	14				
9	Taz6Tbia Rectilinear 60 T(19)	189	Taz6Tbia Rectilinear 60 T(19)	14	52	Taz6Tbia Rectilinear 60 T(52)	31	Taz6Tbia Rectilinear 60 T(63)	15	Taz6Tbia Rectilinear 60 T(63)	15				
10	Taz6Tbia Rectilinear 60 T(20)	136	Taz6Tbia Rectilinear 60 T(20)	45	53	Taz6Tbia Rectilinear 60 T(53)	23	Taz6Tbia Rectilinear 60 T(64)	16	Taz6Tbia Rectilinear 60 T(64)	16				
11	Taz6Tbia Rectilinear 60 T(21)	189	Taz6Tbia Rectilinear 60 T(21)	10	54	Taz6Tbia Rectilinear 60 T(54)	25	Taz6Tbia Rectilinear 60 T(65)	33	Taz6Tbia Rectilinear 60 T(65)	33				
12	Taz6Tbia Rectilinear 60 T(22)	194	Taz6Tbia Rectilinear 60 T(22)	11	55	Taz6Tbia Rectilinear 60 T(55)	49	Taz6Tbia Rectilinear 60 T(66)	10	Taz6Tbia Rectilinear 60 T(66)	10				
13	Taz6Tbia Rectilinear 60 T(23)	178	Taz6Tbia Rectilinear 60 T(23)	72	56	Taz6Tbia Rectilinear 60 T(56)	31	Taz6Tbia Rectilinear 60 T(67)	7	Taz6Tbia Rectilinear 60 T(67)	7				
14	Taz6Tbia Rectilinear 60 T(24)	128	Taz6Tbia Rectilinear 60 T(24)	15	57	Taz6Tbia Rectilinear 60 T(57)	35	Taz6Tbia Rectilinear 60 T(68)	17	Taz6Tbia Rectilinear 60 T(68)	17				
15	Taz6Tbia Rectilinear 60 T(25)	204	Taz6Tbia Rectilinear 60 T(25)	47	58	Taz6Tbia Rectilinear 60 T(58)	43	Taz6Tbia Rectilinear 60 T(69)	15	Taz6Tbia Rectilinear 60 T(69)	15				
16	Taz6Tbia Rectilinear 60 T(26)	203	Taz6Tbia Rectilinear 60 T(26)	14	59	Taz6Tbia Rectilinear 60 T(59)	49	Taz6Tbia Rectilinear 60 T(70)	10	Taz6Tbia Rectilinear 60 T(70)	10				
17	Taz6Tbia Rectilinear 60 T(27)	120	Taz6Tbia Rectilinear 60 T(27)	67	60	Taz6Tbia Rectilinear 60 T(60)	36	Taz6Tbia Rectilinear 60 T(71)	83	Taz6Tbia Rectilinear 60 T(71)	83				
18	Taz6Tbia Rectilinear 60 T(28)	147	Taz6Tbia Rectilinear 60 T(28)	9	61	Taz6Tbia Rectilinear 60 T(61)	32	Taz6Tbia Rectilinear 60 T(72)	68	Taz6Tbia Rectilinear 60 T(72)	68				
19	Taz6Tbia Rectilinear 60 T(29)	71	Taz6Tbia Rectilinear 60 T(29)	10	62	Taz6Tbia Rectilinear 60 T(62)	31	Taz6Tbia Rectilinear 60 T(73)	14	Taz6Tbia Rectilinear 60 T(73)	14				
20	Taz6Tbia Rectilinear 60 T(30)	67	Taz6Tbia Rectilinear 60 T(30)	37	63	Taz6Tbia Rectilinear 60 T(63)	36	Taz6Tbia Rectilinear 60 T(74)	9	Taz6Tbia Rectilinear 60 T(74)	9				
21	Taz6Tbia Rectilinear 60 T(31)	99	Taz6Tbia Rectilinear 60 T(31)	27	64	Taz6Tbia Rectilinear 60 T(64)	42	Taz6Tbia Rectilinear 60 T(75)	10	Taz6Tbia Rectilinear 60 T(75)	10				
22	Taz6Tbia Rectilinear 60 T(32)	99	Taz6Tbia Rectilinear 60 T(32)	12	65	Taz6Tbia Rectilinear 60 T(65)	46	Taz6Tbia Rectilinear 60 T(76)	10	Taz6Tbia Rectilinear 60 T(76)	10				
23	Taz6Tbia Rectilinear 60 T(33)	115	Taz6Tbia Rectilinear 60 T(33)	23	66	Taz6Tbia Rectilinear 60 T(66)	31	Taz6Tbia Rectilinear 60 T(77)	13	Taz6Tbia Rectilinear 60 T(77)	13				
24	Taz6Tbia Rectilinear 60 T(34)	70	Taz6Tbia Rectilinear 60 T(34)	20	67	Taz6Tbia Rectilinear 60 T(67)	19	Taz6Tbia Rectilinear 60 T(78)	9	Taz6Tbia Rectilinear 60 T(78)	9				
25	Taz6Tbia Rectilinear 60 T(35)	100	Taz6Tbia Rectilinear 60 T(35)	11	68	Taz6Tbia Rectilinear 60 T(68)	18	Taz6Tbia Rectilinear 60 T(79)	7	Taz6Tbia Rectilinear 60 T(79)	7				
26	Taz6Tbia Rectilinear 60 T(36)	58	Taz6Tbia Rectilinear 60 T(36)	20	69	Taz6Tbia Rectilinear 60 T(69)	21	Taz6Tbia Rectilinear 60 T(80)	16	Taz6Tbia Rectilinear 60 T(80)	16				
27	Taz6Tbia Rectilinear 60 T(37)	41	Taz6Tbia Rectilinear 60 T(37)	19	70	Taz6Tbia Rectilinear 60 T(70)	34	Taz6Tbia Rectilinear 60 T(81)	11	Taz6Tbia Rectilinear 60 T(81)	11				
28	Taz6Tbia Rectilinear 60 T(38)	49	Taz6Tbia Rectilinear 60 T(38)	14	71	Taz6Tbia Rectilinear 60 T(71)	70	Taz6Tbia Rectilinear 60 T(82)	8	Taz6Tbia Rectilinear 60 T(82)	8				
29	Taz6Tbia Rectilinear 60 T(39)	60	Taz6Tbia Rectilinear 60 T(39)	44	72	Taz6Tbia Rectilinear 60 T(72)	96	Taz6Tbia Rectilinear 60 T(83)	11	Taz6Tbia Rectilinear 60 T(83)	11				
30	Taz6Tbia Rectilinear 60 T(40)	93	Taz6Tbia Rectilinear 60 T(40)	11	73	Taz6Tbia Rectilinear 60 T(73)	46	Taz6Tbia Rectilinear 60 T(84)	36	Taz6Tbia Rectilinear 60 T(84)	36				
31	Taz6Tbia Rectilinear 60 T(41)	78	Taz6Tbia Rectilinear 60 T(41)	34	74	Taz6Tbia Rectilinear 60 T(74)	36	Taz6Tbia Rectilinear 60 T(85)	38	Taz6Tbia Rectilinear 60 T(85)	38				
32	Taz6Tbia Rectilinear 60 T(42)	60	Taz6Tbia Rectilinear 60 T(42)	10	75	Taz6Tbia Rectilinear 60 T(75)	38	Taz6Tbia Rectilinear 60 T(86)	38	Taz6Tbia Rectilinear 60 T(86)	38				
33	Taz6Tbia Rectilinear 60 T(43)	53	Taz6Tbia Rectilinear 60 T(43)	12											

## REFERENCES

- [1] R. Langner, “Stuxnet: Dissecting a Cyberwarfare Weapon,” *IEEE Security Privacy*, vol. 9, no. 3, pp. 49–51, May 2011.
- [2] D. Formby, S. Durbha, and R. Beyah, “Out of control: Ransomware for industrial control systems,”
- [3] C. Fachkha, E. Bou-Harb, A. Keliris, N. Memon, and M. Ahamad, “Internet-scale probing of cps: Inference, characterization and orchestration analysis,” 2017.
- [4] R. Smith, “8 hot 3d printing trends to watch in 2016,” *Forbes*,
- [5] “Hardware meets software in advanced manufacturing,” <https://www.ge.com/stories/hardware-meets-software-advanced-manufacturing>, 2017.
- [6] “Arconic strengthens 3d printing collaboration with airbus,” <http://advancedmanufacturing.org/arconic-airbus-3d-printing-collaboration/>, 2016.
- [7] F. Jeff, “SpaceX unveils its 21st century spaceship,” *NEWSPACE Journal*, May 30, 2014.
- [8] A. Davies, Feb. 28, and. 6, “A swedish automaker is using 3d printing to make the world’s fastest car,” *Business Insider*,
- [9] G. D. Janaki Ram, Y. Yang, and B. E. Stucker, “Effect of process parameters on bond formation during ultrasonic consolidation of aluminum alloy 3003,” *Journal of Manufacturing Systems*, vol. 25, no. 3, pp. 221–238, 2006.
- [10] B. Berman, “3-d printing: The new industrial revolution,” *Business Horizons*, vol. 55, no. 2, pp. 155–162,
- [11] “Natural machines: The makers of foodini - a 3d food printer making all types of fresh, nutritious foods.,” <http://www.naturalmachines.com/>, 2017.
- [12] J. Hicks, “FDA approved 3d printed drug available in the US,” *Forbes*,
- [13] T. Wohlers, *Wohlers report 2015: 3d printing and additive manufacturing state of the industry; annual worldwide progress report*. Wohlers Associates, 2015.



- [14] M. Yampolskiy, A. Skjellum, M. Kretzschmar, R. A. Overfelt, K. R. Sloan, and A. Yasinsac, "Using 3d printers as weapons," *International Journal of Critical Infrastructure Protection*, vol. 14, pp. 58–71, Sep. 2016.
- [15] L Sturm, C Williams, J Camelio, J White, and R Parker, "Cyber-physical vulnerabilities in additive manufacturing systems," *Context*, vol. 7, no. 2014, p. 8, 2014.
- [16] A. Ipakchi and F. Albuyeh, "Grid of the future," *IEEE power and energy magazine*, vol. 7, no. 2, pp. 52–62, 2009.
- [17] A. K.Z.K. Z. Security, *Inside the Cunning, Unprecedented Hack of Ukraines Power Grid*.
- [18] L. Garcia, F. Brasser, M. H. Cintuglu, A.-R. Sadeghi, O. Mohammed, and S. A. Zonouz, "Hey, my malware knows physics! attacking plcs with physical model aware rootkit," in *24th Annual Network & Distributed System Security Symposium (NDSS)*, Feb. 2017.
- [19] S. R. Chhetri, C. Arquimedes, and M. A. A. Faruque, "KCAD: Kinetic Cyber Attack Detection Method for Cyber-Physical Additive Manufacturing Systems," *Proceedings of the 35th International Conference on Computer-Aided Design*, no. 74, 2016.
- [20] C. Song, F. Lin, Z. Ba, K. Ren, C. Zhou, and W. Xu, "My smartphone knows what you print: Exploring smartphone-based side-channel attacks against 3d printers," in *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, ser. CCS '16, New York, NY, USA: ACM, 2016, pp. 895–907, ISBN: 978-1-4503-4139-4.
- [21] A. Hojjati, A. Adhikari, K. Struckmann, E. Chou, T. N. Tho Nguyen, K. Madan, M. S. Winslett, C. A. Gunter, and W. P. King, "Leave your phone at the door: Side channels that reveal factory floor secrets," in *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, ser. CCS '16, New York, NY, USA: ACM, 2016, pp. 883–894, ISBN: 978-1-4503-4139-4.
- [22] P. C. Baker, M. D. Judd, and S. D. J. Mcarthur, "A frequency-based RF partial discharge detector for low-power wireless sensing," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 17, no. 1, pp. 133–140, Feb. 2010.
- [23] A. Nesbitt, B. G. Stewart, S. G. McMeekin, S. Conner, J. C. Gamio, K. Liebech-Lien, H. O. Kristiansen, and S. Krakenes, "A novel approach to high voltage substation surveillance using radio frequency interference measurement," in *2009 IEEE Electrical Insulation Conference*, May 2009, pp. 159–163.

- [24] M. B. Cohen, U. S. Inan, and E. W. Paschal, "Sensitive Broadband ELF/VLF Radio Reception With the AWESOME Instrument," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 1, pp. 3–17, Jan. 2010.
- [25] M. B. Cohen, R. K. Said, and U. S. Inan, "Mitigation of 5060 Hz power line interference in geophysical data," *Radio Science*, vol. 45, no. 6, RS6002, Dec. 2010.
- [26] K. L. Cummins and M. J. Murphy, "An Overview of Lightning Locating Systems: History, Techniques, and Data Uses, With an In-Depth Look at the U.S. NLDN," *IEEE Transactions on Electromagnetic Compatibility*, vol. 51, no. 3, pp. 499–518, Aug. 2009.
- [27] S. B. Moore, W. B. Glisson, and M. Yampolskiy, "Implications of Malicious 3d Printer Firmware," Jan. 2017, ISBN: 978-0-9981331-0-2.
- [28] M. Backes, M. Drmuth, S. Gerling, M. Pinkal, and C. Sporleder, "Acoustic side-channel attacks on printers," in *Proceedings of the 19th USENIX Conference on Security*, ser. USENIX Security'10, Berkeley, CA, USA: USENIX Association, 2010, pp. 20–20.
- [29] Avery Li-Chun Wang, "An industrial strength audio search algorithm," *ISMIR*, 2003.
- [30] D. Will, *Dejavu*; available at <https://github.com/worldveil/dejavu>, 2017.
- [31] U. S.D.o. E. JMesserly SVG version by User:J, *English: Simple diagram of electricity grids in North America*. Dec. 2008.
- [32] R. K. Said, *Accurate and efficient long-range lightning geo-location using a vlf radio atmospheric waveform bank*. Stanford University, 2009.
- [33] "List of extreme points of the United States," *Wikipedia*, Apr. 2017, Page Version ID: 776549077.
- [34] R. Said, U. Inan, and K. Cummins, "Long-range lightning geolocation using a vlf radio atmospheric waveform bank," *Journal of Geophysical Research: Atmospheres*, vol. 115, no. D23, 2010.
- [35] D. Scherrer, M. Cohen, T. Hoeksema, U. Inan, R. Mitchell, and P. Scherrer, "Distributing space weather monitoring instruments and educational materials worldwide for IHY 2007: The AWESOME and SID project," *Advances in Space Research*, vol. 42, no. 11, pp. 1777–1785, Dec. 2008.